

317.471

# Álkalmasított matematikai lapok

12. ~~AA~~

2000/1

20  
2000

A MAGYAR TUDOMÁNYOS AKADÉMIA  
MATEMATIKAI TUDOMÁNYOK  
OSZTÁLYÁNAK KÖZLEMÉNYEI

20.

KÖTET

# ALKALMAZOTT MATEMATIKAI LAPOK

## A MAGYAR TUDOMÁNYOS AKADÉMIA MATEMATIKAI TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

ALAPÍTOTTÁK

KALMÁR LÁSZLÓ, TANDORI KÁROLY, PRÉKOPA ANDRÁS, ARATÓ MÁTYÁS

FŐSZERKESZTŐ

BENCZÚR ANDRÁS

FŐSZERKESZTŐ-HELYETTESEK

DEMETROVICS JÁNOS, FARKAS MIKLÓS

FELELŐS SZERKESZTŐ

SZÁNTAI TAMÁS

A SZERKESZTŐBIZOTTSÁG TAGJAI

Arató Mátyás, Csirik János, Csiszár Imre, Galántai Aurél, Gécseg Ferenc, Gyires Béla, Györfy László, Harnos Zsolt, Hatvani László, Heppes András, Kátai Imre, Katona Gyula, Kis Ottó, Klafszky Emil, Kovács Margit, Lovász László, Maros István, Prékopa András, Recski András, Stoyan Gisbert, Szentkúti Zsolt (technikai szerkesztő), Tandori Károly, Tusnády Gábor, Varga László

XX. kötet 1. szám

Szerkesztőség és kiadóhivatal: 1027 Budapest. Fő u. 68.

Az Alkalmazott Matematikai Lapok változó terjedelmű füzetekben jelenik meg, és olyan eredeti tudományos cikkeket publikál, amelyek a gyakorlatban, vagy más tudományokban közvetlenül felhasználható új matematikai eredményt tartalmaznak, illetve már ismert, de színvonalas matematikai apparátus újszerű és jelentős alkalmazását mutatják be. A folyóirat közöl cikk formájában megírt, új tudományos eredménynek számító programokat, és olyan, külföldi folyóiratban már publikált dolgozatokat, amelyek magyar nyelven történő megjelentetése elősegítheti az elért eredmények minél előbbi, széles körű hazai felhasználását. A szerkesztőbizottság bizonyos időnként lehetővé kívánja tenni, hogy a legjobb cikkek nemzetközi folyóiratok különszámaként angol nyelven is megjelenhessenek.

A folyóirat feladata a Magyar Tudományos Akadémia III. (Matematikai) Osztályának munkájára vonatkozó közlemények, könyvismertetések stb. publikálása is.

A kéziratok a főszerkesztőhöz, vagy a szerkesztőbizottság bármely tagjához beküldhetők. A főszerkesztő címe:

Benczúr András, főszerkesztő  
1027 Budapest, Fő u. 68.

Közlésre el nem fogadott kéziratokat a szerkesztőség lehetőleg visszajuttat a szerzőhöz, de a beküldött kéziratok megőrzéséért vagy továbbításáért felelősséget nem vállal.

Az Alkalmazott Matematikai Lapok előfizetési ára kötetenként 850 forint. Megrendelések a szerkesztőség címén lehetségesek.

A Magyar Tudományos Akadémia III. (Matematikai) Osztálya a következő idegen nyelvű folyóiratokat adja ki:

1. Acta Mathematica Hungarica,
2. Acta Physica Hungarica.

## AZ INFLÁCIÓ ÉS A KÜLSŐ-BELSŐ EGYENSÚLY EGYÜTTMOZGÁSA\*

MELLÁR TAMÁS

Budapest

A tanulmány a rendszerváltás utáni magyar kormányok inflációs politikáinak mechanizmusát kíséri meg feltárni. Ennek keretében azt vizsgálja, hogy vajon a külső és a belső egyensúlyi helyzet javítása érdekében felhasználják-e az infláció mesterséges növelését eszközként. Az empirikus vizsgálatok alapján az állapítható meg, hogy az egyes kormányok nem voltak erősen elkötelezettek az infláció elleni harcban, esetenként feláldozták az infláció visszaszorításának célját. Az infláció cél- és eszközjellegének ciklikus változására vonatkozó jelek ugyan akadtak, de a választási logikának megfelelő négy éves politikai üzleti ciklusok létét nem lehetett egyértelműen igazolni.

### Bevezetés

A rendszerváltás után eltelt tíz esztendő bizonyos területeken már elegendő összehasonlítható adatot teremtett ahhoz, hogy ne csak elvi síkon, hanem gyakorlati alapon is elemezni lehessen bizonyos makro-összefüggések érvényesülését. A munkanélküliség és az infláció közötti sajátos kapcsolatot megfogalmazó *Phillips-görbe* összefüggés érvényesülésének a vizsgálata az átmeneti gazdaságokban annál is inkább érdekes, mert a fejlett piacgazdaságokban való érvényesülése is erősen vitatott.<sup>1</sup>

Az elmúlt évek adatai fényében elég egyértelműen megállapítható, hogy a magyar gazdaságban nem érvényesült a munkanélküliség-infláció átváltás, hanem sokkal inkább az egyirányú együttmozgás volt jellemző a két makro-változóra. A kiváltó ok azonban eltérően a fejlett piacgazdaságoktól, nem a racionális várakozásokban illetve a pozitív vagy negatív kínálati sokkok létében, hanem a külkereskedelem-szimmetrikus ciklusokban keresendő. A kereslet növekedése azért nem okozott áremelkedést, mert az import kínálat megnövekedése le tudta kötni azt, sőt a termelési

---

\*Ez a dolgozat a XVII. Magyar Operációkutatási Konferencián (Veszprém, 1999. október 7–9.) elhangzott előadás írott változata.

<sup>1</sup>Elég itt csak a keynesiánusok és a monetaristák, illetve újklasszikusok között immár két évtizede zajló vitára utalni.

eszköz import gyakran a hazai kínálati potenciált is növelte, csökkentve ezáltal az egységköltségeket. A külkereskedelmi hiány növekedése azonban előbb-utóbb kényszerítette a forint leértékelését és az import visszafogását. A kettő együtt a termelés csökkenését, a munkanélküliség növekedését és az infláció emelkedését idézte elő.<sup>2</sup>

A kormányok számára önmagában az infláció nagysága is fontos tényező, éppen ezért befolyásolása is igen fontos döntési paraméter. Közismert, hogy a növekvő és magas infláció a költségvetés számára *seigniorage-bevételt* és a köz számára átháríthatatlan *inflációs adót* jelent. A másik oldalon a magas infláció költsége az államadósság utáni nagyobb kamatfizetési kötelezettség, a kedvezőtlenebb folyó költségvetési pozíció és az üzleti szféra elbizonytalanodása. A negatív hatások ellenére bizonyos helyzetekben az egyes kormányok mégis az infláció tudatos felpörgetésére fanyalodnak, majd később annak csökkentésére kényszerülnek. Nem voltak kivételek az ilyen típusú viselkedésmódtól a rendszerváltás utáni magyar kormányok sem. Az érvényesülő ciklikus működési mechanizmus azonban némiképpen bonyolultabb a fent leírtaknál. Az infláció mellett szerepet kap még a kormány által tervezett infláció és a gazdasági szereplők inflációs várakozásai, illetve ezek dinamikus mozgásai. A tényleges és a tervezett infláció, valamint a belső és külső egyensúlyi helyzet évenkénti alakulását az 1. táblázat tartalmazza.

1. táblázat: A tervezett és tényleges infláció, valamint az egyensúlyi helyzet alakulása

Év	Tervezett infláció	Tényleges infláció	Tényleges – tervezett infláció	Költségvetési egyenleg/GDP	Külkereskedelmi egyenleg/GDP
1990	19–20	28,9	9	–0,1	0,2
1991	35–38	35,0	–1	–3,0	0,8
1992	22–25	23,0	–1	–7,0	0,9
1993	17–19	22,5	4,5	–6,7	–9,0
1994	18–22	18,8	–1,2	–9,6	–9,4
1995	19–20	28,2	8,7	–7,3	–5,6
1996	19–20	23,6	4,1	–4,6	–3,7
1997	18–19	18,3	–0,2	–4,7	–2,2
1998	13–14	14,3	0,8	–4,8	–4,8
1999	10–11	9–10	–1	–4,5	–5,0

Forrás: KSH évkönyvek és a PM Költségvetési előirányzatai

A konkrét mechanizmus leírásánál azzal a feltételezéssel élünk, hogy a kormány lényegét tekintve képes befolyásolni az infláció alakulását és minden időszakra meg

<sup>2</sup>Lásd erről bővebben Mellár (1997) tanulmányát.



tudja határozni a számára legmegfelelőbb inflációs szintet. A gazdasági szereplők a folyamatok értékelése alapján az inflációs várakozásaikat alakítják ki. A kormányzat számára előnyös, ha a tervezett inflációnál a tényleges infláció valamivel magasabb lesz. Két megszorítást azonban tennünk kell ehhez a megállapításhoz: csak akkor előnyös, ha a várt infláció viszonylag alacsony marad, valamint ha a tényleges infláció sem válik túl magassá (nem éri el azt a szintet, amely után már kontrollálhatatlanná válik). Miért előnyös a kormány számára ez a helyzet? Azért, mert így a költségvetést könnyebben ki tudja egyensúlyozni, a bevételek magasabbak a kiadások pedig alacsonyabbak lesznek, a bértárgyalásokon pedig alacsonyabb reálbérben tudnak megállapodni, hiszen a tervezett inflációt veszik a nominálbér növekedés alapjául és nyilván hasonló megfontolásokból az ellentétes előjelű eltérés a tervezett és a tényleges infláció között viszont kedvezőtlen a kormány számára. A dolog szépséghibája természetesen az, hogy a tervezettnél magasabbra kalibrált infláció óhatatlanul az inflációs várakozások és az infláció emelkedését váltja ki, s ezért az inflációs költségek előbb vagy utóbb túlszárnyalják a költségvetési előnyöket. Az inflációs költségek közül talán a legjelentősebb az államadósság utáni kamatfizetési kötelezettség emelkedése, amely igen erőteljesen rontja a költségvetési egyenleget is. Ezért aztán a kormány rákényszerül az infláció csökkentésére. A csökkenő infláció mellett amikor azonban még magasak az inflációs várakozások és a tervezett infláció, a tervezett és a tényleges infláció közötti különbségből adódó veszteségek már megjelennek, de egy ideig még kompenzálhatók lesznek az infláció mérséklődéséből következő haszonnal. Amikor azonban már az infláció további csökkentése nagy többlettel nem kecsegtet, akkor a kormány ismét elszánja magát a tervezett inflációt meghaladó inflációt generálni. S ezzel eljutottunk a ciklus kiindulópontjára.

A folyamatot leíró makro-modell kereteit és alapvonásait a következő részben tárgyaljuk. Majd a harmadik rész az empirikus elemzést, az adatok alapján történő verifikálást tartalmazza. S végül a negyedik rész a levonható következtetéseket fogalmazza meg.

### Az inflációs mechanizmus alapmodellje

A probléma tanulmányozásához szükséges egyszerű modellnek két szereplője van, az állam és a gazdasági szereplők együttese. A modellben két endogén változó van a *várt infláció* és a *tervezett infláció*, s ugyancsak két egzogén változó szerepel a *költségvetési* és a *külkereskedelmi egyenleg*, valamint egy döntési változó a *tényleges infláció*. A modellt két viselkedési függvény és egy veszteségfüggvény együttese alkotja. Ezek konkrétan a következők:

$$(1) \quad \pi_t^e = \alpha_1 \pi_{t-1} + \alpha_2 (\pi_{t-1} - \pi_{t-1}^P) \quad \alpha_1, \alpha_2 > 0$$

$$(2) \quad \pi_t^P = \beta_1 \pi_{t-1} + \beta_2 D_{t-1} + \beta_3 K_{t-2} \quad \beta_1, \beta_2, \beta_3 > 0$$

$$(3) \quad L_t = \gamma_1(\pi_t^P - \pi_t) + \gamma_2\pi_t + \gamma_3(\pi_t - \pi_t^e) \quad \gamma_1, \gamma_2, \gamma_3 > 0,$$

ahol  $\pi$  az inflációt jelöli, a felső  $e$  index a várható, a  $P$  index pedig a tervezett inflációra utal, az alsó indexek pedig az időt mutatják,  $D$  a költségvetési (államháztartási) egyenleg,  $K$  pedig a külkereskedelmi (folyó fizetési) mérleg egyenleg GDP-hez viszonyított arányát mutatja. Az első egyenlet a gazdasági szereplők inflációs várakozásainak képződési szabályát rögzíti, amely legközelebbi rokonságba talán az adaptív várakozási szabállyal hozható. Ez két részből (pontosabban azok súlyozott összegéből) tevődik össze: egyfelől a korábbi inflációs ráta nagyságából, másfelől pedig az előző időszakai infláció és a tervezett infláció eltéréséből. Jól látható, hogy a gazdasági szereplők az inflációs várakozások növelésével „büntetik” a kormányt a tervezett infláció túllépéséért. A második egyenlet a tervezett infláció meghatározásának módját definiálja. Ennek első része azt rögzíti, hogy a tervezett infláció erősen függ az előző időszakai inflációtól. Az egyenlet második fele pedig azt mutatja, hogy a kormány által tervezett infláció annál kisebb, minél rosszabb a belső és külső egyensúlyi helyzet a megelőző időszakban. A  $D$  és  $K$  egyensúlyi változó aszerint pozitív, illetve negatív, hogy a költségvetési és a külkereskedelmi mérleg szufficitese vagy deficites. A harmadik egyenlet a kormányzat veszteségfüggvénye, amely három részből áll. Az első rész a tervezett és a tényleges infláció eltéréséből származó veszteséget számszerűsíti. Mivel a súlyparaméterek pozitívak, ezért egyértelműen látszik, hogy a tervezettnél kisebb tényleges infláció a kormányzat számára veszteséget jelent, és fordítva: a tervezettnél nagyobb infláció viszont hasznot (negatív veszteséget) jelent. A második rész az inflációból következő általános veszteséget mutatja, amely a szakirodalomból közismert. A harmadik rész pedig a kormányzat reputációs, bizalmi tőkéjének veszteségét jelenti, amelyet a várthoz képest felfelé eltérített infláció fejez ki.

A veszteségfüggvényt a két magatartási függvény segítségével a következő redukált alakra hozhatjuk:

$$(4) \quad L_t = (\gamma_2 + \gamma_3 - \gamma_1)\pi_t + [\gamma_1\beta_1 - \gamma_3(\alpha_1 + \alpha_2)]\pi_{t-1} + \gamma_3\alpha_2\beta_1\pi_{t-2} + \\ + \gamma_1\beta_2D_{t-1} + \gamma_3\alpha_2\beta_2D_{t-2} + \gamma_1\beta_3K_{t-1} + \gamma_3\alpha_2\beta_3K_{t-2}.$$

A könnyebb áttekinthetőség kedvéért vezessük be a  $c_1, c_2, \dots, c_7$  paramétereket, ekkor a (4) veszteségfüggvény a következő formát veszi fel:

$$(4a) \quad L_t = c_1\pi_t + c_2\pi_{t-1} + c_3\pi_{t-2} + c_4D_{t-1} + c_5D_{t-2} + c_6K_{t-1} + \\ + c_7K_{t-2} \quad c_1, c_2 \geq 0, \quad c_3, \dots, c_7 > 0.$$

A kormányzat feladata tehát úgy meghatározni az aktuális,  $t$ -edik időszakai inflációt,  $\pi_t$ -t, hogy a veszteség minimális legyen. A feladat ilyen formájú meghatározása impliciten feltételezi a kormánynak azt a képességét, hogy jól tudja kontrollálni az inflációt a monetáris és a fiskális politikán keresztül. Ez a gyakorlatban természetesen nem teljesül maradéktalanul, mert a belső piaci erők és a külpia

hatások igen lényeges befolyással bírnak az infláció alakulására. De ezzel együtt is bizonyos határok között a központi intézkedések (tarifák, adózási szabályok, bér- és jövedelem-megállapodások) változtatása révén a kormánynak lehetősége van rövidtávon befolyásolni az árszínvonal-változását. A rendszerváltás utáni magyar gyakorlat és az ezt rögzítő árstatisztika egyértelműen visszaigazolja ezt: a piaci indíttatású ármozgások mellett mindig markáns szerepet játszott a központi intézkedések által generált árváltozások hatása az infláció alakulására. Az ilyen típusú eltérések lényegében a költségvetési politika eszközeivel valósulnak meg, s ezért joggal merülhet fel, hogy vajon a monetáris politika miért nem képes ellensúlyozni a fiskális indíttatású inflációgerjesztő akciókat? Nem a jegybanki függetlenség, vagy az akarat hiánya miatt nem vállalkozik erre az MNB, hanem azért mert a pénzügyi intézményrendszer még nem elég fejlett és a monetáris eszközök nem elég erősek az önálló inflációs célkitűzés, illetve az aktív antiinflációs beavatkozások megvalósításához.<sup>3</sup>

A redukált formában felírt célfüggvényből egyértelműen kiláglik, hogy a kormányzat számára tényleges döntési paraméter csak az aktuális infláció és nem a tervezett infláció. A tervezett infláció meghatározása egy magatartási szabály alapján történik, amely ugyan a megelőző időszak infláció mellett az egyensúlyi helyzet értékelését is magában foglalja, de még ezzel együtt sem tekinthető optimalizálandó döntési paraméternek. Egyrészt azért nem, mert a tervezett inflációnak mindig a tárgyidőszakot megelőzően már meghatározottnak kell lennie, hiszen ennek alapján történnek a bértárgyalások és a következő évi költségvetési számok kialakítása. Másfelől pedig, a tervezett inflációnak elég szoros összhangban kell lennie az általános inflációs folyamattal, a korábbi időszakok inflációs adataival, annak érdekében, hogy egy elfogadható szinten hiteles lehessen a gazdasági szereplők előtt.

Visszatérve a veszteségfüggvényünk elemzésére, mindenek előtt azt a sajátosságot kell megmagyaráznunk, ami a  $c_4, \dots, c_7$  paraméterek pozitivitásából következik. Ez ugyanis azt jelenti, hogy az egyensúlyi helyzet javulása (pozitív, szufficites mérleg) növeli a veszteségeket, rontja a kormányzat pozícióját. Az ellentmondás azonban csak látszólagos, amely abból következik, hogy a veszteségfüggvény az inflációra van felírva. A kedvező egyensúlyi helyzet ( $D, K > 0$ ) folyamánként ugyanis nem lesz feszített az inflációs terv, vagyis viszonylag magas lesz a tervezett infláció, illetve várhatóan magasabb lesz, mint a tényleges infláció (mert a rossz egyensúlyi helyzet nem készíti a kormányt a tudatos alultervezésre). Ezért aztán veszteség éri a kormányt, mert a költségvetési kiadásai a magasabb tervezett inflációhoz igazodnak, a bevételei pedig az alacsonyabb tényleges inflációhoz.

A veszteségfüggvény egyszerűbb alakra hozása és további elemzése céljából az eddig egzogenként kezelt egyensúlyi mutatókat kell „endogenizálni”. Nyilván mind a belső, mind a külső egyensúly jelentős mértékben függ az infláció nagyságától.

<sup>3</sup> Jól példázza ezt a csúszóleértékelési rendszer működtetése, amely igen fontos a külső pénzügyi egyensúly és a bizalom megtartása szempontjából, de éppen ezért az árfolyampolitika nem használható antiinflációs célok megvalósítására. A monetáris intézmények és eszközök gyengeségéről a három átmeneti országban további adalékok olvashatók Brada-Kutan (1999) tanulmányában.

Tételezzük fel, hogy a következő egyszerű kapcsolat áll fenn az egyensúlyi változók és az infláció között:<sup>4</sup>

$$D_t = \omega \pi_t \quad \omega \geq 0$$

$$K_t = \phi \pi_t \quad \phi < 0.$$

Az  $\omega$  paraméter lehet pozitív és negatív is, annak megfelelően, hogy a seigniorage és inflációs adó bevétel nagyobb, vagy kisebb, mint az adósságszolgálat növekedése. A  $\phi$  paraméter egyértelműen negatív, mert az infláció következtében romlik a versenyképesség és ez a külkereskedelmi mérleg romlásához vezet. Ezt a két összefüggést behelyettesítve a (2) egyenletbe a következőképpen módosul a tervezett inflációra vonatkozó összefüggés:

$$(2a) \quad \pi_t^P = \beta \pi_{t-1} \quad \text{ahol} \quad \beta = \beta_1 + \beta_2 \omega + \beta_3 \phi$$

Az előzőekből következően a  $\beta$  paraméter lehet pozitív és negatív egyaránt. A (2a)-t figyelembe véve a veszteségfüggvény redukált alakja a következő lesz:

$$(4b) \quad L_t = c_1 \pi_t + c_2 \pi_{t-1} + c_3 \pi_{t-2} \quad c_1, c_2, c_3 \geq 0$$

A  $c$  értékek itt csak annyiban különböznek az előbbi (4a) formától, hogy most a meghatározatlan előjelű  $\beta$  helyettesíti a pozitív  $\beta_1$ -et.

Amennyiben a (4b) veszteségfüggvény elemzése során az inflációról feltételezzük, hogy nem lehet negatív (jöllehet ez általánosságban erős kikötés, de a rendszer-váltás utáni magyar helyzetet alapul véve egyáltalán nem az), vagyis  $\pi_t \geq 0$  minden  $t$ -re, akkor a következő megfontolások tehetők a veszteségfüggvény minimalizálása kapcsán.

- Ha a  $c$  paraméterek mindegyike pozitív, akkor az infláció optimális nagysága nulla minden időszakra, s a kumulált veszteség is nulla lesz, bármilyen időhorizont vonatkozásában.
- Ha  $c_1, c_2, c_3 < 0$ , akkor az optimális megoldás a minél nagyobb infláció választása minden időszakra, mert ez minimalizálja a veszteségfüggvényt. Ebben az esetben a rendszer teljesen egyensúlytalanná válik, explodál.
- Minden más esetben „vegyes megoldás” lesz az optimális, vagyis speciális keveréke a pozitív és a zero inflációs időszakoknak. Jól látható ez a veszteségfüggvény véges,  $n$ -időszakra összesített formájából:

$$\sum_{i=1}^n L_i = (c_1 + c_2 + c_3)[\pi_1 + \pi_2 + \dots + \pi_{n-2}] + (c_1 + c_2)\pi_{n-1} + c_1 \pi_n$$

<sup>4</sup>A definiált kapcsolat több szempontból is erősen leegyszerűsített: egyrészt azért, mert csak a jelenbeni inflációtól való függést kezeli, másrészt azért, mert a közvetett hatásokat nem veszi figyelembe (hiszen az egyensúlyi változók maguk is hatást gyakorolnak az inflációra), harmadrészt pedig azért, mert az egyéb, nem inflációs tényezőket figyelmen kívül hagyja.

Annak megfelelően, hogy a  $(c_1 + c_2 + c_3)$ , valamint a  $(c_1 + c_2)$  paraméter-együttesek, illetve a  $c_1$  paraméter pozitív vagy negatív előjelű, lesz az érintett időszakos infláció zéró, vagy pozitív.

Jól látható, tehát, hogy az előre meghatározatlan előjelű paramétereknek kiemelt jelentőségük van a teljes folyamat alakulásában. Ez persze nem véletlen, hiszen a paraméterek előjel-irányának konkrét közgazdasági jelentősége van. Ha ugyanis  $c_1 > 0$  ez azt jelenti, hogy  $\gamma_2 + \gamma_3 > \gamma_1$ , ami pedig arra utal, hogy a kormány számára nagyobb jelentőségűek az infláció növekedéséből származó költségek, mint a tervezettől elmaradó tényleges infláció okozta veszteségek. Vagyis  $c_1$  pozitivitása egy, az infláció ellen komolyan elkötelezett kormány jelenlétére, negativitása pedig egy valószínűsítően nem elkötelezett, az inflációt más célok érdekében feláldozható eszközként kezelő kormányt feltételez. A  $c_2 > 0$  azt jelzi, hogy  $\gamma_1\beta > \gamma_3(\alpha_1 + \alpha_2)$ , ami úgy interpretálható, hogy az előző időszakokból áthúzódó inflációs hatás negatív következményeit a kormány nagyobbra értékeli, mint a pozitív következményeket. S végül a  $c_3 = \gamma_3\alpha_2\beta > 0$  azt mutatja, hogy a kormány hajlik arra, hogy a korábbi infláció alapján a tervezett inflációt magasabbra állítsa be, s így kevésbé legyen lehetősége az infláció felpörgetésével haszonra szert tennie.

### Empirikus vizsgálatok

Az adatokon alapuló vizsgálat három fő részre bontható: elsőként a (4a) összefüggés, az eredeti modell redukált formájának a tesztelése történik, majd a leegyszerűsített, endogenizált egyensúlyi változók mellett származtatott (4b) egyenlet tesztelése következik, és végül az (1-2), illetve az (1-2a) egyenletekből származtatott inflációs várakozási magatartási szabály empirikus igazolása zárja az elemzést.

A makrogazdasági működés sajátos reakcióideiből következően negyedéves léptékű adatok elemzése látszik célszerűnek. Rendelkezésünkre állnak 1989 és 1998 között az előző év azonos negyedéven alapuló negyedéves fogyasztói árindexek, viszonylag homogén módszertani alapon számítva. A szokásoknak megfelelően mi is az éves (12 havi) intervallumú fogyasztói árindexet tekintjük az infláció mérőszámának, tehát a továbbiakban ezzel dolgozunk, illetve ennek szezonálisan kiigazított változatával. A költségvetési deficit és a külkereskedelmi mérlegre vonatkozóan csak az 1992 és 1998 közötti időszakra vannak összehasonlítható negyedéves adataink, ami némiképpen korlátozza a tesztelés időintervallumát.

### Az infláció és a külső-belső egyensúlyi helyzet kapcsolata

Mindenek előtt célszerű elvégezni a változóinkra az egységgyök tesztet, annak érdekében hogy megállapítsuk, milyen integráltságú idősoraink vannak. A kibővített *Dickey-Fuller* egységgyök teszt elvégzése alapján megállapítható, hogy a magyar inflációs folyamat és az egyensúlyi helyzetet jellemző költségvetési valamint külkereskedelmi egyenleg elsőrendű integráltságú (lásd a 2. táblázatot)

2. táblázat: Az inflációs és egyensúlyi idősorok egységgyök teszt eredményei

	Infláció	Várt infláció	Költségvetési egyenleg	Külkereskedelmi mérleg
Szint	-2,569 (3) -1,014 (4)	-2,75 (3) -1,45 (4)	-1,378 (3) -2,100 (2)	-2,82 (2) -2,63 (1)
Első differencia	-3,839*** (3) -2,827* (4)	-3,562** (3) -3,01** (4)	-3,767*** (3) -6,708*** (2)	-3,757*** (2) -4,42*** (1)

\* Szignifikáns 10%-os szinten, \*\* szignifikáns 5%-os szinten, \*\*\* szignifikáns 1%-os szinten

A zárójelben az optimális késleltetés értéke, amelyet az Akaike információs kritérium és a Schwartz kritérium segítségével határoztunk meg. A számításokat minden esetben konstans mellett végeztük.

A (4a) összefüggés teszteléséhez célszerűnek látszik először egy kicsit átrendezni az összefüggést. Ha feltesszük, hogy a veszteségfüggvény több időszakon keresztül egy jellemző (optimális?) szinten van, akkor  $L_t = b$  helyettesítéssel élhetünk, s ennek megfelelően a következő egyszerűsített formát kapjuk:

$$(5) \quad \pi_t = a_0 + a_1\pi_{t-1} + a_2\pi_{t-2} + a_3D_{t-1} + a_4D_{t-2} + a_5K_{t-1} + a_6K_{t-2}$$

ahol  $a_0 = b/c_1$ ,  $a_1 = -c_2/c_1$ , ...,  $a_6 = -c_7/c_1$ . Mivel a változók közötti kapcsolatok vonatkozásában nem zárható ki elméletileg a fordított hatás sem, nevezetesen az infláció befolyása az egyensúlyi változókra, ezért kézenfekvőnek tűnik egy VAR-modell felállítása az érintett három változó, az infláció, a költségvetési és a külkereskedelmi egyenleg bevonásával.<sup>5</sup> A becslési eredményeket a 3. táblázat tartalmazza.

A becslési eredményekből megállapítható, hogy csak az inflációra felírt összefüggés fogadható el, a külkereskedelmi mérleg esetén egyedül a saját kettővel késleltetett értéke szignifikáns, a költségvetési egyenleg tekintetében pedig egyik tényező sem. A kettős késleltetés nem csak elméleti szempontból volt kíváncsú, hanem gyakorlati oldalról is elfogadhatónak bizonyult, a likelihood-arány teszt bizonyította megfelelőségét. A becslési eredmények alapján felírható az (5) becslt alakja:

$$(5a) \quad \pi_t = 1,309\pi_{t-1} - 0,405\pi_{t-2} - 1,734D_{t-1} + 0,48D_{t-2} + 0,43K_{t-1} - 0,675K_{t-2}$$

Mivel a konstans nem volt szignifikáns, ezért már a specifikáció során kihagytuk. Közgazdasági oldalról ez azt jelenti, hogy a veszteségfüggvény jellemző értéke a vizsgált időintervallumban nulla volt, vagyis a rendszerváltás utáni kormányok igyekeztek minimalizálni a veszteségeiket. Ha feltételezzük, hogy a meghatározatlan előjelű  $c_1$  paraméter pozitív (egyébként ezt támasztja alá a becslt célfüggvény-érték

<sup>5</sup>Mivel a változók nem kointegráltak, ezért egységgyökök jelenléte mellett is alkalmazható a VAR-modell és konzisztens becslést ad.



3. táblázat. A VAR-modell eredményei

	$\pi_t$	$K_t$	$D_t$
$\pi_{t-1}$	1,309 (7,78)	-0,056 (-0,55)	0,028 (0,34)
$\pi_{t-2}$	-0,405 (-2,66)	0,054 (0,58)	-0,055 (-0,74)
$K_{t-1}$	0,429 (1,31)	0,145 (0,73)	0,048 (0,29)
$K_{t-2}$	-0,675 (-2,20)	0,416 (2,23)	0,191 (1,27)
$D_{t-1}$	-1,714 (-3,51)	0,518 (1,76)	0,031 (0,13)
$D_{t-2}$	0,480 (0,86)	0,186 (0,56)	0,046 (0,17)
$R^2$	0,891	0,306	-0,095
$F$ -statisztika	32,92	1,76	-0,34
Log likelihood	-47,92	-34,80	-29,38

Zárójelben a  $t$ -próba értéke

is, vagyis az a feltételezés, hogy a kormányok általában nagyobbra értékelik az infláció költségét, mint a belőle származó előnyöket), akkor a becült paraméterek többségének előjele megfelel az elméleti elvárásoknak. Kivételt csak az  $a_4, a_5 > 0$  paraméterértékek jelentenek, ezek azonban nem szignifikánsak, ahogy az a 3. táblázatban közölt  $t$ -értékekből kiolvasható. A szignifikáns egyensúlyi változók negatív hatása a modell logikájából következően azt jelenti, hogy a javuló egyensúlyi helyzet által csökkenti az inflációt, hogy mérsékli a kormányzat infláció felpörgetési szándékát (tervezett fölé emelését, mert a tervezett infláció is reálisan és nem alá van tervezve). Természetesen nem zárható ki a modell logikájától független, a standard közgazdasági elméleten alapuló értelmezés sem, amelynek értelmében a költségvetési deficit mérséklődése következtében csökken a kereslettöbblet, a külkereskedelmi egyenleg javulása pedig enyhíti a leértékelésre irányuló nyomást, s ezek miatt csökken az infláció.

### A hazai inflációs folyamat jellege

A (4b) redukált veszteségfüggvény tesztelése tulajdonképpen a magyar inflációs folyamat dinamikájának empirikus vizsgálatát jelenti. Az előző részhez hasonlóan itt is egyszerűbb formára hozhatjuk az összefüggést a konstans célfüggvény-érték feltételezésével. Ennek megfelelően a következő egyenletet kapjuk:

$$(6) \quad \pi_t = a_0 + a_1 \pi_{t-1} + a_2 \pi_{t-2},$$

ahol  $a_0 = b/c_1$ ,  $a_1 = -c_2/c_1$ ,  $a_6 = -c_3/c_1$ . A (6) differenciaegyenlet az ökonometriai vizsgálatok szemszögéből egy autoregresszív folyamatnak tekinthető, amelynek érvényesülése az adatok alapján közvetlenül tesztelhető. Több ARMA-moddal való kísérletezés után azt találtuk, hogy az inflációs idősorhoz egy AR(2) folyamat igen jól illeszthető. A becült adatok alapján tehát a (6) egyenlet konkrét formája a következő (zárójelben a *t*-próba értékek):

$$(6a) \quad \pi_t = 22,356 + 1,4063\pi_{t-1} - 0,53\pi_{t-2}$$

$$(7,1) \quad (9,7) \quad (-3,6)$$

$$R^2 = 0,86$$

$$D.-W. = 2,18$$

$$F\text{-stat.} = 111,37$$

A differenciaegyenlet gyökei könnyen számíthatók, ezek  $\lambda_1 = 0,70 + 0,19i$  és  $\lambda_2 = 0,70 - 0,19i$ . Jól látható, hogy komplex gyököket kaptunk, ezért ennek megfelelően a (6a) általános megoldása a kezdőértékek függvényében a következő:

$$(6b) \quad \pi_t = [0,73^t] [(\pi_0 - 179,3) \cos 15^\circ t + (5,3\pi_1 - 3,7\pi_0 - 283,1) \sin 15^\circ t] + 179,3$$

A  $\pi_0$  és a  $\pi_1$  kezdőértékek megadásával bármely időpontra kiszámítható az aktuális inflációs szint. A paraméterek értelmezésénél nézzük először a pozitív konstans tagot! Ez azt jelenti, hogy hosszabb idő átlagában a kormányzat (illetve több kormány is) nem törekedett a zéró veszteség elérésére, hanem engedett a rövidtávú csábításnak, amelynek következtében hosszabb távon társadalmi veszteség alakult ki. Ez a megállapítás ellentmond az (5) és (5a) egyenlet teszteredményeinek. Viszont ha az előző résznek megfelelően most is feltesszük, hogy a  $c_1 > 0$ , akkor a becült paraméterértékek alapján a  $c_2 < 0$  és  $c_3 > 0$  adódik, ami teljesen egybevág az előző rész eredményeivel. Mindazonáltal a külső és belső egyensúlyi helyzet endogenizálása némileg módosította a kormányzatok gazdaságpolitikájáról levonható következtetéseket. Még az egzogén egyensúlyi mutatók melletti becslésnél elég egyértelműen egy infláció ellen elkötelezett magatartást lehetett kimutatni, addig most kevésbé elkötelezett magyar kormányok képe jelenik meg. A pozitív veszteségfüggvény-érték és az infláció ciklikus pályája azt jelzi, hogy itt könnyen előfordulhat, hogy esetenként az infláció felpörgetésének eszközához nyúlnak. Alapvetően azonban nem kaptunk más természetű viselkedést, hiszen az most is igaz, hogy a kormányok preferencia rendszerükben az infláció költségeit nagyobbra értékelték, mint a belőle származó előnyöket. Ezt a megállapítást támasztja alá a (6b) általános megoldásból következtető néhány megállapítás is. Mivel a konjugált komplex gyökök négyzetgyöke (0,728) kisebb, mint 1, ezért az inflációs folyamat stabil, tart a 179,27-es egyensúlyi szintjéhez.

A kiegészítő megoldás második szögletes zárójelen belüli részében megjelenő trigonometrikus függvényből az következik, hogy az időbeli pálya fluktuáló és periodikus. Egy periódus hossza 24 időszak (360/15), vagyis 6 esztendő, minthogy negyedéves léptékű időszakokkal dolgozunk. Ez utóbbi nem igazolja vissza a negy éves választási periódusok kézenfekvő szakaszolási idejét, amely a kiinduló koncep-

ciónkban benne rejlik. Ugyanis elég nyilvánvaló a választások utáni első két éves időszakot, mint az egyensúlyteremtés időszakát tekinteni, tehát ilyenkor lódul meg az infláció, s a választások előtti két évet pedig, mint a költségvetési engedés, vagyis az alacsony infláció időszakát tekinteni. Az egész inflációs folyamatba azonban zavaró tényezőként számítható be a rendszerváltás körüli időszak, ahol a különféle sokkok következtében ugrásszerűen emelkedett meg az infláció szintje. Ha ezt az időszakot elhagyjuk, s a becslést csak az 1992–99-es időszakra végezzük el, akkor a következő eredményt kapjuk:

$$\begin{aligned}
 (6c) \quad \pi_t &= 19,96 + 1,28\pi_{t-1} - 0,47\pi_{t-2} \\
 &\quad (7,8) \quad (7,2) \quad (-2,8) \\
 R^2 &= 0,79 \quad D.-W. = 1,95 \quad F = 49,23
 \end{aligned}$$

Ebben az esetben a komplex gyökök értéke  $\lambda = 0,64 \pm 0,24i$ , amiből a 4,3 év periódusidő adódik.<sup>6</sup> Ez igazolhatja a fentebb megfogalmazott politikai üzleti ciklusok létre vonatkozó hipotézisünket. Azonban nagyon óvatosan kell kezelni az itt kapott eredményeket, mert egyfelől a modell az egyensúlyi változók olyan endogenizálásán alapul, amely elméletileg nem eléggé megalapozott. Másrészt, mert a becsült paraméterekből számított periódushossz szerfelett bizonytalan (még a konfidencia intervallumon belül maradó paraméterérték módosulás is jelentősen megváltoztathatja az egész folyamat jellegét, pl. nem is ciklikus lesz a folyamat). Harmadrészt pedig, mert az empirikus elemzés alapjául szolgáló minta túl kicsi (csak két periódust fog át és ebből nem nagyon lehet stabil viselkedési szabályokra következtetni).

### Az inflációs várakozások számszerűsítése

A további empirikus vizsgálatok számára még hátra van az inflációs várakozásokra és a tervezett inflációra vonatkozó (1)–(2), illetve az (1)–(2a) egyenletek tesztelése. A tervezett inflációra vonatkozó (2) illetve (2a) összefüggéseket behelyettesítve a várt infláció azt kapjuk, hogy

$$\begin{aligned}
 (7) \quad \pi_t^e &= (\alpha_1 + \alpha_2)\pi_{t-1} - \alpha_2\beta_1\pi_{t-2} - \alpha_2\beta_2D_{t-2} - \alpha_2\beta_3K_{t-2} \\
 \pi_t^e &= a_1\pi_{t-1} + a_2\beta_1\pi_{t-2} + a_3D_{t-2} + a_4K_{t-2} \quad a_1 > 0, \quad a_2, a_3, a_4 < 0
 \end{aligned}$$

illetve

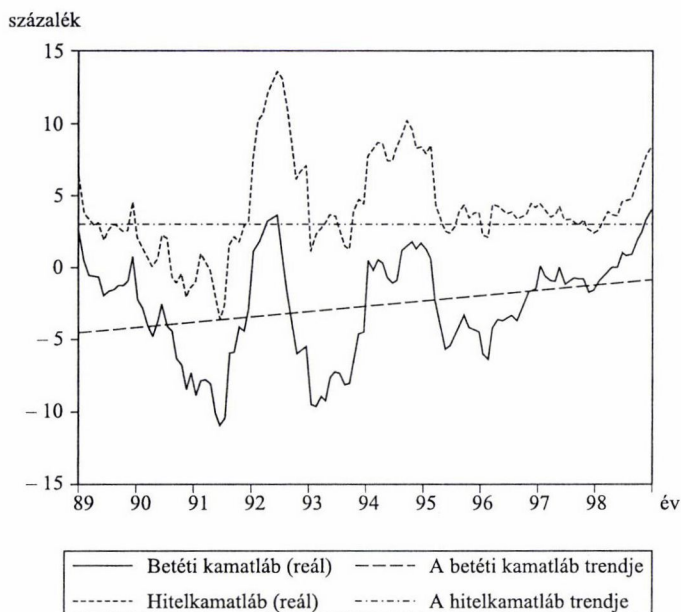
$$\begin{aligned}
 (7a) \quad \pi_t^e &= (\alpha_1 + \alpha_2)\pi_{t-1} - \alpha_2\beta\pi_{t-2} \\
 \pi_t^e &= a_1\pi_{t-1} + a_2\pi_{t-2} \quad a_1 > 0, \quad a_2 < 0
 \end{aligned}$$

Ezt a két összefüggést kell tehát tesztelni a következőkben. Az inflációs várakozások ökonometriai vizsgálatánál az alapvető problémát az jelenti, hogy a várt inflációra

<sup>6</sup>Mert  $\cos \alpha = 0,64/\sqrt{0,64 + 0,24} = 0,9363$ , ami  $20^\circ 30'$ -nek felel meg és  $360/20,5 = 4,3$ .

vonatkozó adatok nem állnak rendelkezésre. Mindenek előtt tehát ezt az akadályt kellett elhárítani a várt infláció meghatározásával.<sup>7</sup>

A várt infláció meghatározásához a reálkamatok alakulását hívtuk segítségül. Ha az éven belüli futamidejű nominális betéti és hitelkamatlábakból a 12 havi inflációs ráta segítségével reálkamatlábát számítunk, akkor azt találjuk, hogy ezek az értékek igen nagy volatilitást mutatnak (lásd az 1. ábrát).



1. ábra. A rövidtávú kamatlábak és a hosszútávú hozamok alakulása

A Fisher azonosság értelmében a reálkamatláb hosszabb időszak átlagában viszonylag állandó és a nominális kamatláb az inflációs várakozások szerint alakul, képletben:

$$R_t = \hat{r}_t + \pi_t^e$$

ahol  $R_t$  a nominális kamatláb,  $\hat{r}_t$  pedig az időben viszonylag állandó reálkamatláb. Mivel a tényleges reálkamatlábakhoz úgy jutottunk, hogy a nominális kamatlábból levontuk a tényleges inflációt

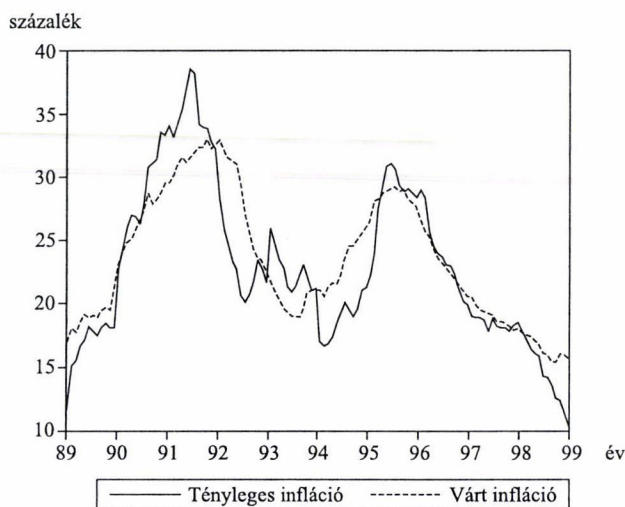
$$r_t = R_t - \pi_t,$$

<sup>7</sup>Közismert ugyan a szakirodalomban az adaptív, illetve a racionális várakozások becslési technikája, de ez megkerüli az inflációs várakozások konkrét meghatározását, ami viszont a mi elemzésünknek lényegi eleme.

ezért egyszerűen származtathatjuk az inflációs várakozásokra a következő összefüggést:

$$(8) \quad \pi_t^e = \pi_t + r_t - \hat{r}_t$$

Ha tehát ismernénk a hosszabb távon érvényesülő reálhozamot, akkor a (8) összefüggés segítségével meg tudnánk határozni a várt infláció nagyságát. Nyilván nem járunk messze az igazságtól, ha az állandó reálkamatlábát egy átlagként, egy trendértékként fogjuk fel. Ennek megfelelően az időben változó reálkamatlábához illesztettünk egy trendet, amelynek értékei a hosszabb távon érvényesülő reálhozam tendenciáját jelzik. Így már lehetővé vált az inflációs várakozások meghatározása. A (8) összefüggés úgy is értelmezhető, hogy a tényleges reálkamatlábak azért és aszerint térnek el a hosszabb távon érvényesülő reálhozam-követelménytől amiért és amennyiben az inflációs várakozás nem volt megfelelő és eltért a tényleges inflációtól. A számításokat elvégezve a betéti és a hitel-kamatlábakra, a várt inflációt a kettő egyszerű számtani átlagaként határoztuk meg.<sup>8</sup> A 2. ábra mutatja a várt és a tényleges infláció alakulását a vizsgált időszakban.



2. ábra. A tényleges és a várt infláció alakulása

Most már abban a helyzetben vagyunk, hogy az összefüggéseinket tesztelhetjük. Vegyük először az egyszerűbb (7a) összefüggést, amely ebben a leegyszerűsített formájában az adaptív várakozási szabályt testesíti meg, hiszen a jelenre vonatkozó inflációs várakozásokat végső soron a korábbi tényleges inflációk súlyozott értéke

<sup>8</sup> A számítások havi kamatlábak alapján történtek, s ennek megfelelően havi gyakoriságú (12 hónapra vonatkozó) inflációs várakozásokat kaptunk, amelyekből egyszerű számtani átlagolással kaptuk a negyedéves indexeket.

határozza meg. Mindenek előtt célszerű megvizsgálni a várt infláció integráltsági fokát. Az egységgyök teszt eredményeit a 2. táblázat tartalmazza. A teszt eredményeiből jól látható, hogy az inflációhoz hasonlóan a várt infláció is egyes integráltságú folyamat. Ebből következően viszont direkt módon, regressziós egyenlet illesztésével nem végezhetjük a tesztelést. Viszont semmilyen elvi akadálya nincs a két változó közötti kointegrációs vizsgálatnak. A *Johansen* kointegrációs teszt eredményei azt mutatják, hogy legfeljebb csak egy kointegrációs kapcsolat meglétét lehet valószínűsíteni a számított likelihood-arányok alapján 5%-os szignifikancia szint mellett. A kointegrációs együtthatók alapján számított hibakorrekciós modell eredménye a következő lett:

$$(8a) \quad \Delta\pi_t^e = -0,026[\pi_{t-1}^e - 1,017\pi_{t-1}] + 0,704\Delta\pi_{t-1}^e + 0,050\Delta\pi_{t-1}$$

(−0,40)                      (−34,7)                      (5,33)                      (0,58)

$$R^2 = 0,589 \qquad F\text{-stat.} = 25,1 \qquad \text{Log likelihood} = -55,7$$

A becslési eredmények azt mutatják, hogy a kointegrációs kapcsolat által meghatározott egyensúlyi szinthez a konvergencia igen bizonytalan, mert a vonatkozó paraméter nem szignifikáns. S ugyancsak nem szignifikáns a múltbeli infláció változásának a hatása sem, amely pedig igen fontos lenne a (7a) érvényességének fennállásához. Az eredmények alapján az is felvethető, hogy egyáltalán van-e valamilyen oksági kapcsolat a várt és a tényleges infláció között. A *Granger*-féle páros oksági vizsgálat eredményét a 4. táblázat foglalja össze.

4. táblázat. A *Granger* oksági vizsgálat eredményei

*A null hipotézis elfogadásának valószínűsége*

Null hipotézis	1 késle- tetés	2 késle- tetés	3 késle- tetés	4 késle- tetés	5 késle- tetés
$\pi$ nem oka $\pi^e$ -nek	0,284	0,585	0,550	0,715	0,184
$\pi^e$ nem oka $\pi$ -nek	0,461	0,106	0,091	0,021	0,015
$\Delta\pi$ nem oka $\Delta\pi^e$ -nek	0,596	0,804	0,740	0,274	0,260
$\Delta\pi^e$ nem oka $\Delta\pi$ -nek	0,500	0,654	0,535	0,015	0,048

Az oksági vizsgálat eredménye eléggé egyértelműen arra utal, hogy az inflációt nem tekinthetjük a várt infláció meghatározó okának, míg a fordított oksági kapcsolat sokkal nagyobb valószínűséggel állítható. Közgazdaságilag ez azt jelenti, hogy az adaptív típusú várakozások nem tekinthetők tipikusnak az inflációs várakozások vonatkozásában. Ugyanakkor viszont a tényleges infláció alakulását lényegesen befolyásolja az inflációs várakozások változása, vagyis nem kizárt, hogy érvényesül a várakozásokkal kibővített Phillips-görbe összefüggés a magyar gazdaságban is.

Az adaptív várakozások hipotézisének elvetése után érdemes ellenpróbaként azt megvizsgálni, hogy vajon a racionális várakozások érvényesültek-e az infláció vonatkozásában. Ezt könnyen megtehetjük, hiszen a racionális várakozások definíciója



értelmében ezek a várakozások nem tartalmazhatnak szisztematikus előrejelzési hibát, vagyis érvényesülnie kell a következő összefüggéseknek:

$$(9) \quad E(\pi_t^e - \pi_t) = 0 \quad \text{és} \quad \pi_t^e - \pi_t = \varepsilon_t,$$

ahol  $\varepsilon_t$  normális eloszlású konstans szórású véletlen változó. Ennek alapján csak azt kell tesztelni, hogy a két változó különbsége vajon tényleg nulla várható értékű normális eloszlású-e. Az erre vonatkozó normalitási teszt eredményei azt mutatják, hogy a *Jarque-Bera* érték 0,45 és a normalitásra vonatkozó null-hipótesis fennállásának a valószínűsége 0,79. Tehát itt is megerősítést kaptunk az adaptív várakozás feltételezésének elvetésére.

Mindez azonban nem jelenti azt, hogy az (1)–(2) egyenletek és a belőlük redukált (7) összefüggés nem igazolható empirikusan, hiszen ott nem tiszta adaptív várakozásokról van szó, mert a (2) egyenlet tartalmazza az egyensúlyi viszonyok változását is, s a racionálisan viselkedő gazdasági szereplők ezt is figyelembe veszik inflációs várakozásaik kialakítása során. Ezért célszerűnek látszik az előző hibakorrekciós modellünket kibővíteni egzogen tényezőként a költségvetési és a külkereskedelmi egyenleg adataival. Több késleltetést végigpróbálva a következő eredmény adódott:

$$(8) \quad \Delta\pi_t^e = 0,101[\pi_{t-1}^e - 1,046\pi_{t-1}] + 0,627\Delta\pi_{t-1}^e + 0,041\Delta\pi_{t-1} - \\ (1,46) \quad (-51,2) \quad (5,53) \quad (0,46) \\ - 0,236\Delta D_{t-1} - 0,233\Delta K_{t-2} \\ (-1,32) \quad (-1,78)$$

$$R^2 = 0,713$$

$$F\text{-stat.} = 12,4$$

$$\text{Log likelihood} = -27,5$$

Az egyensúlyi helyzet alakulásának a beiktatása hozott némi pozitív változást az eredményekben. Egyrészt javult a becslés minősége, jöllehet az előző hiányosságok egy része továbbra is fennmaradt. A hosszú távú egyensúlyhoz való konvergencia nem valósul meg, sőt itt még a mozgás iránya sem megfelelő, mert a konvergencia-paraméter pozitív, s az infláció változása továbbra sem szignifikáns. A belső és a külső egyensúlyi helyzet paramétereinek az előjele megfelel az elméleti várakozásoknak, bár szignifikanciájuk nem elég meggyőző. Mindezek alapján csak fenntartásokkal fogadhatjuk el a (8) becslést az (1)–(2) modell, illetve a (7) összefüggés igazolásának.

### Összegzés

A tanulmányban megfogalmazott modell és annak empirikus verifikálása alapján elég nagy biztonsággal állítható, hogy Magyarországon a rendszerváltás utáni

kormányok nem voltak egyértelműen elkötelezettek az inflációellenes küzdelemben. Az általuk alkalmazott gazdaságpolitikák gyakran nem célváltozóként, hanem az egyensúlyi helyzet javításához szükséges eszközváltozóként kezelték. A cél- és eszközjelleg dominanciájának időbeli megoszlására az ökonometria vizsgálatok nem adtak (az alkalmazott modell jellegéből következően nem is adhattak) kellő felvilágosítást, így e vonatkozásban csak arra az alapadatok által nem cáfolt feltételezésre alapozhatunk, hogy a nagy külső és belső egyensúlytalanság és ezért erőteljes egyensúlyteremtés időszakában az infláció jelentősen emelkedett, tehát főként eszközként szerepelt, míg a kiegyensúlyozott időszakokban célként szerepelt, vagyis lényegesen alacsonyabb értéket vett fel.

A negyedéves inflációs adatokon becsült autoregresszív modell nem támasztja alá meggyőzően a hazai inflációs folyamat ciklikus jellegére vonatkozó hipotézist. A választási időtávnak megfelelő 4 év körüli periódusidő nem bizonyítható, a periódusidő hossza igen bizonytalan, erőteljesen függ a becslési időtartam hosszának megválasztásától. A becslési eredmények alapján ugyan valószínűsíthető a politikai üzleti ciklus logikája szerint cselekvő kormányok jelenléte, de az idő rövidsége és az adatok hiányossága miatt nem dokumentálható és nem bizonyítható megfelelő szinten.

A reálkamatlábakból származtatott inflációs várakozásokkal végzett vizsgálat megmutatta, hogy az adaptív várakozások feltételezése a magyar gazdaságban nem állja meg a helyét. Sokkal inkább tekinthető az inflációs várakozások képződése racionális jellegűnek a rendszerváltás utáni időszakban. Ezt támasztja alá az is, hogy pusztán a korábbi időszakok inflációs számai nem megfelelően magyarázzák a várható infláció alakulását, ha azonban a korábbi makro-egyensúlyi helyzet jelzőszámaait is figyelembe vesszük, akkor már többé-kevésbé elfogadható becslést kapunk.

## Irodalom

- [1] A. Alesina, Politics and business cycles in Industrial Democracies, *Economic Policy*, (4) (1989), 57–98.
- [2] A. Banerjee and D. Hendry, *The econometrics of economic policy*, Blackwell Publishers (Oxford, UK, 1997).
- [3] J. C. Brada and A. M. Kutan, *The end of moderate inflation in three transition economies?* Arizona State University and Federal Reserve Bank of St. Louis. (1999).
- [4] W. W. Charemza and D. F. Deadman, *New directions in econometric practice*. Edward Elgar Publishing Co. (Cheltenham, UK, 1997).
- [5] Csermely Ágnes, Az inflációs célkitűzés rendszere, *Közgazdasági Szemle*, 3 (1997), 233–253.
- [6] R. F. Engle and C. W. J. Granger, Co-integration and Error Correction: Representation, Estimation and Testing, *Econometrica*, 2 (1987), 251–276.
- [7] Ékes Ildikó, Az infláció okai Magyarországon, *Statisztikai Szemle*, 4–5 (1998), 358–374.
- [8] G. Gandolfo, *Economic dynamics*, Springer Verlag (Berlin–Heidelberg, 1997).
- [9] J. Hamilton, *Time series analysis*, Princeton University Press (1994).

- [10] D. Hendry, *Dynamic econometrics*, Oxford University Press (1995).
- [11] D. Hendry, Econometric Modelling with Cointegrated Variables: An Overview, *Oxford Bulletin of Economics and Statistics*, 2 (1986), 201–212.
- [12] Hunyadi László, Egységgyökök és tesztjeik, *Sigma*, 3 (1994), 135–164.
- [13] Kovács Erzsébet, Idősorok kointegrációja, *Statisztikai Szemle*, 6 (1989), 599–619.
- [14] Kőrösi Gábor, Mátyás László és Székely István, Gyakorlati ökonometria, *KJK, Budapest*, (1990), 481.
- [15] Lieli Róbert, Az idősormodelleken alapuló inflációs előrejelzések I.-II.: egyváltozós módszerek és többváltozós módszerek, *MNB Füzetek*, 4 (1999) és kézirat.
- [16] Mellár Tamás, Makrogazdasági célok, gazdaságpolitikai ciklusok, *Sigma*, 4 (1997).
- [17] Simon András, Az infláció tényezői 1990–95-ben, *MNB Füzetek 1.*, Budapest, (1996), 10.
- [18] S. Turnovsky, *Methods of macroeconomic dynamics*, The MIT Press (Cambridge Massachusetts, 1995).

(Beérkezett: 1999. szeptember 20.)

MELLÁR TAMÁS  
KÖZPONTI STATISZTIKAI HIVATAL  
1024 BUDAPEST, KELETI K. U. 5-7.  
E-mail: TAMAS.MELLAR@KSH.X400GW.ITB.HU

## PARALLEL MOVEMENT OF INFLATION, DOMESTIC AND FOREIGN EQUILIBRIA

T. MELLÁR

The paper attempts to reveal the mechanism of the inflationary policy of the Hungarian governments of the transition period. It investigates whether forced inflation is used as a means of improvement of foreign and domestic equilibria. According to the empirical results it can be deduced that the governments were not highly committed in the struggle against inflation, sometimes they even sacrificed the goal of suppressing inflation. There are, however, some indications, for a cyclical alteration of the pattern of inflation but the existence of the four year business cycles corresponding to those of the general elections could not be unequivocally proven.



## NÉHÁNY POLINOMIÁLIS PROGRAMOZÁSI FELADAT MEGOLDÁSA KLASZTEREZŐ ALGORITMUSSEL\*

BORGULYA ISTVÁN\*\*

Pécs

Tanulmányomban egy olyan klaszterező algoritmust ismertetek, amely lineáris, vagy nemlineáris feltételrendszerek figyelembevételével, egyaránt alkalmazható multimodális függvények adott számú lokális szélsőértékének, vagy globális optimumának közelítő meghatározására. Az algoritmus a feltételeket a célfüggvénytől elkülönítve kezeli, és az evolúciós algoritmusoknál alkalmazott egyik mérték (measure of violation of constraints) alapján veszi figyelembe őket. Az algoritmus a szokásos tesztfeladatok esetén kb. 0.01 pontossággal határozza meg a szélsőértékek helyét és értékét. A lehetőségek szemléltetésére, a tesztfeladatok mellett, néhány polinomiális programozási feladat megoldását ismertetem.

### 1. Bevezetés

Számos technikai és gazdasági probléma megoldásához olyan nemlineáris függvényt kell optimalizálni, melynek több lokális minimuma (maximuma) található a vizsgált tartományban. E multimodális függvényeknél általában a globális optimumot keressük, de egyes esetekben szükséges a lokális szélsőértékek ismerete is. E tanulmányban egy újabb módszert kívánok e szélsőérték keresésre bemutatni.

A vizsgált probléma matematikai formája a következő: keressük az  $f$  multimodális függvény egy  $x^* \in S$  globális minimumpontját, amelyre

$$x^* = \min \left( \{ f(x) \mid x \in F \subseteq S \} \right).$$

Legyen  $S$  egy  $m$ -dimenziós tartomány:

$$S = \{ x \in \mathbb{R}^m \mid a_i \leq x_i \leq b_i, \ a_i, b_i, x_i \in \mathbb{R}, \ i = 1, 2, \dots, m \},$$

---

\*Ez a dolgozat a XVII. Magyar Operációkutatási Konferencián (Veszprém, 1999. október 7–9.) elhangzott előadás írott változata.

\*\*Készült az OTKA T030861 kutatás keretében.

ahol az  $F \subseteq S$  tartományt  $r$  ( $r \geq 0$ ) számú lineáris és/vagy nemlineáris feltétellel adunk meg:

$$g_j(x) \leq 0, \quad j = 1, 2, \dots, r.$$

A feltételrendszerbeli egyenlőségeket, pl. a  $h(x) = 0$ -t, helyettesítsük a

$$-h(x) - \delta \leq 0,$$

$$+h(x) - \delta \leq 0,$$

párral, ahol a  $\delta$  paraméter az algoritmus pontosságát fogja jelenteni.

Nézzük először a rendelkezésre álló módszereket. A korlátozási feltétel nélküli, globális optimalizációs problémát megoldó algoritmusok két nagy csoportba sorolhatók: garantált megbízhatóságú, valamint heurisztikus módszerek csoportjára (pl. [11], [17]-ben). A garantált megbízhatóságú módszerek közt a legfontosabb az intervallum felosztási módszerek csoportja. E módszerek adott pontossággal képesek megtalálni a globális optimumot, esetleg exponenciálisan növekvő lépésszám mellett. Alapgondolatuk az, hogy a tartományt egyre kisebb részekre osztják és a megoldás szempontjából egyes részeket elvetnek, vagy előnyben részesítenek. A részleteket tekintve különböző változatok léteznek (pl. [16], [14], [25]). E módszerek általában csak akkor használhatók, ha a függvény megfelelő analitikai tulajdonságokkal rendelkezik.

Ellentétben az előző csoporttal a heurisztikus módszerek csak nagy valószínűséggel határozzák meg a globális optimumot. Alkalmazásukhoz elegendő ismerni a függvény értékeit az adott pontokban, a függvény analitikai tulajdonságait nem használják fel. E csoport legismertebb módszerei a sztochasztikus módszerek, melyek véletlen mintapontok alapján keresik a lokális minimumokat, valamint a simulated annealing és evolúciós technikák, melyek fizikai, ill. biológiai analógiák alapján közelítenek a globális optimumhoz ([6], [18], [1], [23]). Nézzük a módszereink szempontjából lényeges sztochasztikus módszereket közelebbről.

A sztochasztikus módszerek körébe különböző véletlen keresési technikák, valamint klaszterező algoritmusok tartoznak (pl. [6]-ban). Egy jelenetős csoportjukat képezik a kétfázisú módszerek, melyeknél az algoritmus működésében két fázis különíthető el. Az első fázisban az algoritmus véletlen mintapontot, ill. mintapontokat generál, majd a második fázisban valamelyik pontból „helyi kereső eljárást” indít egy-egy lokális minimum megkeresésére. Ilyen módszer pl. a multistart, amely minden generált pontból helyi kereső eljárást indít, vagy a klaszterező algoritmusok (pl. [30], [31], [5], [32]-ben). Klaszterezésnél folyamatosan csoportosítják a generált mintapontokat, és minden csoportnál helyi kereső eljárást indítanak egy lokális minimum meghatározására. E kétfázisú módszerek alkalmasak minden lokális minimum keresésére.

További lényeges csoportja a sztochasztikus módszereknek a random search módszerek csoportja. E módszerek valamilyen valószínűségi eloszlás szerint generált mintapontok alapján keresik a globális optimumot (pl. [28]).



A különböző feltételrendszerrel adott optimalizációs problémák megoldására szintén számos módszer alkalmazható (lásd pl. [2]). E módszerek arra használhatók, hogy a problémát egy ekvivalens, korlátozási feltétel nélküli, vagy egy egyszerű feltételrendszerű globális optimalizációs problémává alakítsuk át. Ehhez pl. a penalty és a barrier függvények, vagy Lagrange duális módszere alkalmazható. Több módszer a lineáris és nemlineáris feltételek esetén más-más megoldási alternatívát ajánl (pl. [33], [34]), azon módszerek pedig, melyek evolúciós algoritmusokat alkalmaznak a feltételrendszerrel adott problémák megoldására, általában szintén a penalty függvény módszerét alkalmazzák ([3]).

A bemutatásra kerülő módszer szintén sztochasztikus módszer, egy klaszterező algoritmus, amely multimodális függvények minimumait (maximumait) határozza meg. Ellentétben az említett klaszterező algoritmusokkal e módszer először kialakítja a klaszttereket, majd utána pontosítja a szélsőértékeket. Az *optimumkereső klaszterező algoritmus* (OSCA) a mintapontokat feldolgozva, folyamatosan úgy alakítja a klaszttereket, hogy a prototípusaik (magpontjaik) lehetőleg minél közelebb kerüljenek egy-egy lokális minimumhoz. A klaszttereket hasonlóság alapján alakítja ki és egy prototípus környezettel (box) „szeparálja” a szélső értékeket egymástól. A [8]-beli algoritmus [9]-beli továbbfejlesztésével az OSCA feltételrendszerrel adott problémák megoldására is felhasználható. A módosítás lényege:  $S$  minden  $x$  pontját az evolúciós algoritmusoknál alkalmazott mértékkel, az  $F$  tartománytól való  $D(x)$  „távolsággal” jellemez, mely távolság alapja a feltételrendszer nemteljesülésének foka (degree of violation of constraints). E  $D(x)$  „távolság” alapján az algoritmus a klasztterek kialakításánál preferál, vagy elvet pontokat, és akár  $F$ -en kívüli pontokat is eredményesen fel tud használni ([3], [22]). Ezzel a célfüggvény és a feltételek kezelését különválasztja.

Összehasonlítva az OSCA, valamint a szokásos klasztterező algoritmusok alap gondolatát, a következő főbb jellemzők emelhetők ki.

- Az OSCA nem az összes, hanem csak  $n$  darab lokális minimumot (maximumot) keres.
- A klasztterek prototípusát (magpontok) a lokális minimum (maximum) pontokkal azonosnak választják a módszerek. A szokásos módszerek e prototípust egy helyi kereső eljárás egyszeri alkalmazásával határozzák meg, míg az OSCA párhuzamosan, folyamatosan javítva közelíti a lokális minimumokat.
- A szokásos módszereknél a csoportosítás alapja a távolság, és minden klaszter kiterjedésére egy „kritikus” távolságot adnak meg. Az OSCA hasonlóság alapján csoportosít és egy elemet a leghasonlóbb prototípus klaszterébe sorol.
- Az OSCA a kritikus távolság helyett egy prototípus környezetet vezet be, melynek felhasználásával a véletlenszerűen kialakított klasztterek közül az azonos szélsőértéket meghatározó klaszttereket időnként kiszűri és csak egyet hagy meg közülük.
- A szokásos módszerek a megállási szabályt különböző szempontok, pl. lokális minimumok strukturája, vagy a függvényértékek eloszlása alapján választják

meg. Az OSCA-nál a prototípusok stabilitása a megállási feltétel. Ha stabilak az értékek, a prototípusok környezetében koncentrált kereséssel még tovább pontosítja a szélsőértékek értékét, helyét, mielőtt megáll.

- Feltételrendszerrel adott optimalizációs problémák esetén az  $F$ -től való „távolság” alapján  $S$  minden pontját fel tudja használni a klaszterek kialakításánál.

Nézzük a továbbiakban először az új klaszterező algoritmust, a teszt eredményeket, majd a polinomiális programozási feladatok megoldását.

## 2. Az OSCA

### 2.1. A kiindulásként választott algoritmus.

A lokális minimumok meghatározásához egy fuzzy klaszterező algoritmust módosítottam ([7]). E kiindulásul választott fuzzy klaszterező algoritmus alap gondolata a következő:

Tekintsük a klaszterezésre kerülő elemeket egy többkritériumos döntési probléma alternatíváinak, és válasszunk egy fuzzy rendező módszert (FRM). Ha minden elemhez (alternatívához) az FRM-el, mint függvénnyel egy értéket rendelünk, ellenőrizhető az elemek sorrendje (FRM értéke függjön mind a kritérium értékektől, mind a kritérium súlyoktól).

Használjuk fel e függvény értékeket a klaszterek kialakításánál. Soroljuk az elemeket hasonlóság alapján  $t$  darab klaszterbe. Rendeljük minden klaszterhez egy centrumot és prototípust. Legyen a centrum a klaszterbeli elemek függvényértékeinek átlaga, a prototípus pedig a klaszter olyan eleme, melynek függvényértéke a legközelebb esik a centrum értékéhez. Az algoritmus a kritériumok súlysúlyszámának változtatásával megkeresi, megtanulja az egyes klasztereket jellemző prototípusokat, centrumokat.

Az osztályozást fokozatosan tanulja meg rendszer. Egyenként, véletlenszerűen veszi elő az elemeket, és minden elemet hasonlósági mérték alapján valamelyik klaszterbe sorol. Hogy a centrumok egy stabil érték felé konvergáljanak, a kritériumok súlysúlyszámait úgy módosítja, hogy az elem új függvényértéke a megfelelő centrum értéke felé közeledjék. A súlysúlyszámokat minden olyan esetben módosítja, ha az utolsó két elemhez tartozó függvényértékek sorrendje eltérő klaszter centrumaik sorrendjétől (A súlysúlyszám módosítás a neurális hálók delta szabályához hasonló képlettel történik). Az algoritmus mindaddig folytatódik, amíg a centrumok stabilá nem válnak (változásuk értéke adott küszöbszám alá nem csökken). (Az algoritmus tehát a klaszterezéshez szükséges ismereteket implicit formában, a klasztereket jellemző prototípusokban és a kritériumok közös súlyaiban tárolja.)

Az OSCA alap gondolata abban tér el e fuzzy klaszterező algoritmus alap gondolatától, hogy a klaszterek prototípusait a lokális minimum pontokkal lehetőleg azonosnak választja és az FRM alkalmazása helyett a multimodális függvény érté-

keit használja fel. Ezek az eltérések lényegesen egyszerűsítik az új algoritmust, így nincs szükség fuzzy technikára és súlyszám tanulásra sem.

A kiindulási algoritmushoz képest a következő eltéréseket kell még az optimalizáláshoz megvalósítani:

1. Mivel szélsőértékeket keresünk, mindig a pillanatnyilag ismert legkisebb (ill. legnagyobb) függvényértékű elemet kell egy klaszter elemei közül kiválasztani. Legyen a klaszter centruma azonos ezen elem függvényértékével, prototípusa pedig a hozzá tartozó input adatokkal.
2. Nem kell a klaszterek elemeit nyilvántartani. Elegendő csak egyetlen elem, a prototípus nyilvántartása.
3. A klaszterek generálása közben ki kell szűrni azon prototípusokat (és klasztereket), melyek egy másik, olyan prototípus környezetébe esnek, melynek centruma kisebb (ill. nagyobb) értékű. Ezt az ellenőrzést elegendő pl. az iteráció minden  $kn$ -dik lépésében végrehajtani.
4. A tesztfeladatok eredményei szerint, az algoritmus pontosabb eredményeket ad, ha az elemekkel a  $[0, 1]$  intervallumra (ill. a  $[0, 1]^m$  hiperkockába) transzformálva számol.

## 2.2. Az algoritmus

OSCA-t, a fenti megállapításokat figyelembe véve, a fuzzy klaszterező algoritmus módosításával kapjuk. Jelöljük a klaszterezésre kerülő  $S$ ,  $m$ -dimenziós tartomány elemeit (alternatíváit)  $a_1, a_2, \dots, a_i, \dots$ -vel, az elemeket leíró kritériumokat  $K_1, K_2, K_m$ -mel, egy  $a_i$  elemet pedig adjunk meg a  $k_{i1}, k_{i2}, \dots, k_{im}$  adatokkal. Az  $a_i$  elemhez tartozó függvényértéket jelöljük  $o_i$ -vel ( $o_i = f(a_i)$ , ahol  $f$  a multimodális függvényt jelöli). Tegyük fel, hogy  $t$  számú lokális minimumot akarunk keresni. (Amennyiben lokális maximumokat keresünk reláció cseréjével módosíthatjuk a kívánt módon az algoritmust).

Vezessük be a következő jelöléseket még:

- Egy  $a_i$  elem környezetét jelölje  $\text{Box}(a_i)$ , amely minden kritérium esetén egyforma, a  $h$  változóval meghatározott szélességű intervallumot jelent:  $\text{Box}(a_i) = \{[k_{ij} - h, k_{ij} + h], j = 1, 2, \dots, m\}$ .
- Két  $a_i, a_j$  elemnél jelölje a hasonlóságot

$$H(a_i, a_j) = 1 / (1 + d(a_i, a_j))$$

ahol  $d(a_i, a_j)$  az elemek euklideszi távolsága.

- Legyen  $D(a_i)$  az  $a_i$  elem „távolsága”  $F$ -től

$$D(a_i) = \left( \sum_{j=1}^r \max \{g_j(a_i), 0\}^2 \right)^{1/2}$$

(Ha az  $a_i \in F$ , vagy nincs feltételrendszer, akkor  $D(a_i) = 0$ ).

Használjuk fel  $D(a_i)$  értékét a pontok jellemzésére a következő módon:

$a_i$  jobb  $a_j$ -nél, ha  $D(a_i) < D(a_j)$ .

$D(a_i) = D(a_j)$  esetén  $a_i$  akkor jobb  $a_j$ -nél, ha  $f(a_i) < f(a_j)$ .

E jellemzők felhasználása lehetővé teszi, hogy az OSCA azon pontokkal is számoljon, melyek nem elemei  $F$ -nek.

Az algoritmus működését hét paraméter:  $t$ ,  $h$ ,  $epsz$ ,  $itt$ ,  $delta$ ,  $kn$  és  $\delta$  befolyásolja. Nézzük külön-külön a szerepüket:

- $t$  – a keresett lokális minimumok számát adja meg.
- $h$  – a prototípus környezetek közös paramétere, amely a Box intervallumainak nagyságát befolyásolja. (Az algoritmusnak a  $[0, 1]$  intervallumra transzformált értéket kell megadni).
- $eps$  – a stabilitási feltétel paramétere. Ha az utolsó két ellenőrzés során a szélsőértékek ingadozása kisebb mint  $eps$  (*error*, *oerror* értékek), stabilnak tekintjük a klasztereket.
- $itt$  – a koncentrált keresés paramétere. Ha az iterációk  $it$  száma eléri az  $itt$  értéket, elkezdődik a prototípusok egyre szűkülő környezetében a koncentrált keresés.
- $delta$  – a koncentrált keresés megállási feltételének paramétere. Koncentrált keresés esetén minden  $kn$ -dik iteráció után a  $h$  értékét felére csökkenti az algoritmus. Az eljárás akkor fejeződik be, ha a  $h$  értéke kisebb mint  $delta$ .
- $kn$  – az ellenőrzések időpontját meghatározó paraméter. A stabilitást csak minden  $kn$ -dik iterációban ellenőrzi az algoritmus.
- $\delta$  – paraméter, amely a feltételrendszerbeli egyenlőségek pontosságát, és egyben az algoritmus pontosságát adja meg.

Az algoritmus lépései a következők:

## 1. Kezdőértékek beállítása.

Adjuk meg  $t$ ,  $h$ ,  $eps$ ,  $itt$ ,  $delta$ ,  $kn$  és  $\delta$  értékét. Legyenek  $oc_1, oc_2, \dots, oc_t$  és  $p_1, p_2, \dots, p_t$  a klaszterek centrumai és prototípusai. Legyen az  $i$ -edik klaszter centruma, jelzőszáma és prototípusa  $oc_i$ ,  $c_i$  és  $p_i$ . Legyen  $it = 0$ ,  $error = 1$ ,  $oerror = 1$ ,  $tuning = 0$ ,  $c_i = 0$ ,  $p_i = 0$ ,  $oc_i = 0$  és  $uoc_i = 0$  ( $i = 1, 2, \dots, t$ ).

## 2. Új elem.

### 2.1. Elem választás.

Ha  $tuning = 0$  generáljunk véletlenszerűen egy  $S$ -beli  $a_j$  elemet.

Ha  $tuning = 1$  generáljunk véletlenszerűen valamely  $Box(p_i)$ -ben egy  $a_j$  elemet.

2.2. Transzformálás. Transzformáljuk  $a_j$ -t a  $[0, 1]$  intervallumba, vagy a  $[0, 1]^m$  hiperkockába.

3. Hasonlóság ellenőrzése.

Legyen  $H(a_j, p_z) = \max_q H(a_j, p_q)$ ;  $q, z \in \{1, 2, \dots, t\}$ .

4. Prototípus módosítás.

Ha  $((o_j < oc_z) \wedge (D(a_j) = D(p_z)) \vee (D(a_j) < D(p_z))) \wedge (c_z = 1)$ , akkor  $p_z = a_j$ ,  $oc_z = o_j$ ,  $cx_z = D(a_j)$ .

5. Új klaszter.

Ha  $(p_z \neq a_j) \wedge (\exists i)(c_i = 0)$  akkor  $p_i = a_j$ ,  $oc_i = o_j$ ,  $c_i = 1$ ,  $cx_i = D(a_j)$ .

6. A centrumok stabilitásának ellenőrzése:

$it = it + 1$ . Ha  $\text{mod}(it, kn) \neq 0$ , folytatása a 2. pontnál, különben:

6.1. Klaszter törlés: minden  $i, ii$  indexre  $(i, ii \in \{1, 2, \dots, t\} \ i \neq ii)$

Ha  $(c_i > 0) \wedge (c_{ii} > 0) \wedge (p_{ii} \in \text{Box}(p_i)) \wedge ((oc_{ii} > oc_i) \wedge (cx_i = cx_{ii}))$   
 $\vee (cx_{ii} > cx_i))$  akkor  $c_{ii} = 0$ ,  $oc_{ii} = 0$ ,  $uoc_{ii} = 0$ .

6.2. Megállási feltétel. Legyen

$$\text{error} = \sum_{i=1}^t |oc_i - uoc_i|^2$$

és  $uoc_i = oc_i$  ( $i = 1, 2, \dots, t$ ).

$\text{tuning} = 0$  esetén: ha  $[(\text{error} < \text{eps}) \wedge (oerror < \text{eps}) \wedge ((\nexists i)(c_i = 0)) \vee (it > itt)]$  hamis,  $oerror = \text{error}$  és folytatás a 2. pontnál. Különben legyen  $\text{tuning} = 1$ ,  $h = h/2$  és folytatás a 2. pontnál.

$\text{tuning} = 1$  esetén: ha  $h < \text{delta}$  igaz, vége az eljárásnak. Különben legyen  $h = h/2$  és folytatás a 2. pontnál.

### A paraméterek megválasztása

A paraméterek közül az  $\text{eps}$ ,  $\text{delta}$ ,  $kn$  értékeire megadható egy standard érték, melynél az algoritmus helyes eredményeket adhat (pl.  $\text{eps} = 0.001$ ,  $\text{delta} = 0.001$ ,  $kn = 500$ ). Ettől kissé eltérő értékeknél az eredmény pontossága, vagy az iterációk száma változhat. Az  $itt$  értéke már erősen feladatfüggő. (Általában 2500 iteráció után már elkezdődhet a koncentrált keresés).

A  $t$  és  $h$  paraméter választása alapvetően befolyásolja az eredményt. Ha lokális minimumokat keresünk, először is meg kell becsülnünk a prototípusok közti legkisebb  $dmin$  távolságot. Legyen  $nm$  a minimumok száma. Ekkor a lehetőségek:

- ha  $t \leq nm$ , a  $h$  értéke határozza meg az eredményt. Ha  $h \in (0.25 \cdot dmin, 0.75 \cdot dmin)$ , helyes eredményt kapunk. Minél nagyobb  $0.75 \cdot dmin$ -nél a  $h$ , annál nagyobb a valószínűsége, hogy nem teljesül a stabilitási feltétel és kevesebb szélsőértéket tud csak megtalálni az algoritmus. Minél kisebb  $0.25 \cdot dmin$ -nél

a  $h$ , annál nagyobb a valószínűsége, hogy nem csak szélsőérték pontok lesznek az eredményben.

- ha  $t > nm$ , ( $n < \infty$ ) nem fog teljesülni a stabilitási feltétel.

Mindkét esetben, ha nem teljesül a stabilitási feltétele, az iteráció csak akkor áll le, ha az iterációk száma nagyobb mint *itt*. A koncentrált keresés után az eredmények közül csak azon pontokat tekinthetjük helyes megoldásnak, melyek értéke több (pl. 1–2 ezer) iteráció óta változatlan.

### 2.3. Teszt eredmények

Bár az OSCA alapvetően multimodális függvények lokális szélsőértékeinek meghatározását teszi lehetővé, alkalmas globális optimum keresésére is. Egy multimodális függvény esetén ugyanis, ha meghatározzuk a minimumhelyeket, egyúttal a legkisebb minimumot, a globális optimumot is megkapjuk. Ez a megoldás azonban nem a leggazdaságosabb. Felhasználhatjuk az OSCA-t a globális optimum keresésére anélkül, hogy minden lokális minimumot megkeresnénk. Ha csak néhány (pl. két) minimumpontot keresünk, koncentrált kereséssel jó közelítést kaphatjuk a globális optimumnak. Unimodális függvények esetén szintén alkalmazható az előbbi módszer. Bár ebben az esetben csak egy szélsőérték létezik, alkalmazható az algoritmus megfelelő *itt* érték válsztásával. Egyszerűbb függvények esetén pontos közelítést kapjuk a globális minimumnak.

Az OSCA működését szemléltesse több ismert mintapélda: alkalmazzuk lokális szélsőértékek, valamint globális optimum keresésére. A példák paraméter beállításait, a futtatások iterációs számát, az eredmények pontosságát az 1. táblázat tartalmazza. (Tekintettel arra, hogy benchmark-ként alkalmazott függvényekről van szó, képleteik, optimális helyeik ismertetésétől eltekintünk). Mivel sztochasztikus módszerről van szó, minden példát ötször futtattunk ismételten, és az *it* iterációs számok (és egyben függvényhívás számok), valamint a pontosság értékek az ismételt futtatások átlag eredményeit mutatják.

a) *Lokális szélsőértékek keresése.* A bemutatásra kerülő tesztfeladatok közül az első, a módosított „Himmelblau” függvény, melynek 4 globális maximuma van, a második pedig a „Shekel’s Foxholes” függvény, 25 lokális minimummal. A két függvény szélsőértékeit OSCA hasonló pontossággal határozta meg mint pl. az evolúciós technikát alkalmazó módszerek (pl. [15], [24]).

b) *Globális optimum keresése.* A különböző globális optimum meghatározó módszerek tesztelésénél a feladatok többsége multimodális függvény, melynek csak a globális minimumát, ill. minimumait keressük, a további lokális szélsőértékek érdektelenek. (Így az OSCA lehetőségeit csak részben kell kihasználni e problémakörnél). Az irodalomban számos tesztfeladat található. A legismertebb Dixon és Szegő (1978) klasszikus tesztfeladat sora, valamint De Jong 5 tesztfüggvénye, de e mellett több összeállítás ismert. Alkalmazzuk a továbbiakban az OSCA bemutatására a klasszikus tesztfeladatokat (jelölés: S5, S7, S10, H3, H6, BR, GP, C6, SCH),



és De Jong függvényeit (jelölés: Sphere, Rosenbrock, step, De Jong F4, Shekel's foxh).

E feladatoknál egységesen két minimumot kerestünk a lehetséges lokális minimumok közül. A paraméter értékek megválasztásában az *itt* és *delta* értékek bizonyultak a legkritikusabbnak. Választásuk nagyban befolyásolta az eredmény pontosságát. A kapott eredmények pontossága általában jobb 0,01-nél, amely elfogadható közelítésnek tekinthető.

	<i>dim</i>	<i>itt</i>	<i>t</i>	<i>delta</i>	<i>it</i>	<b>pontosság</b>
<b>Himmelblau</b>	2	2000	4	$10^{-3}$	4491	$10^{-2}$
<b>Shekel's foxh.</b>	2	13000	25	$10^{-4}$	18463	$10^{-2}$
<b>S5</b>	4	1500	2	$10^{-3}$	3992	$10^{-2}$
<b>S7</b>	4	1500	2	$10^{-3}$	3992	$10^{-2}$
<b>S10</b>	4	1500	2	$10^{-3}$	3992	$10^{-2}$
<b>H3</b>	3	400	2	0.005	1497	$10^{-3}$
<b>H6</b>	6	400	2	0.005	1497	$10^{-4}$
<b>BR</b>	2	400	2	0.005	1497	$10^{-2}$
<b>GP</b>	2	400	2	0.005	1497	$10^{-4}$
<b>C6</b>	2	400	2	0.005	1497	$10^{-3}$
<b>SCH</b>	2	400	2	0.005	1497	$10^{-2}$
<b>Sphere</b>	30	2500	2	$10^{-3}$	5489	$10^{-3}$
<b>Rosenbrock</b>	2	1500	2	0.005	2994	$10^{-5}$
<b>Step</b>	5	1500	2	0.05	2994	$10^{-5}$
<b>De Jong F4</b>	30	2500	2	$10^{-3}$	6487	$10^{-5}$
<b>Shekel's foxh.</b>	2	1500	2	0.005	2994	$10^{-3}$

1. táblázat. Paraméterek és futási eredmények ( $\epsilon ps = 0,001$ ,  $kn = 500$ ,  $h = 0,1$ )

A klasszikus tesztfeladatok eredményeit összehasonlítva más módszerekkel, főleg a sztochasztikus módszereknel kapunk hasonló iterációkat (függvény hívásokat). Nézzük a szokásos, függvényhívások száma alapján való összehasonlítást. Több módszer, és köztük az OSCA eredményét a 2. táblázat tartalmazza ([12], [5], [11], [19], [4], [29] alapján). Megállapítható, hogy az OSCA leginkább Törn és Price módszerével vethető egybe. Az OSCA egyike azon módszereknek, melyek mind a 9 tesztfeladatot sikeresen oldják meg.

c) *Feltételrendszerrel adott globális optimum keresése.* Az OSCA a  $D(a_i)$  „távolság” alkalmazásával lényegesen leegyszerűsíti a feltételrendszerek kezelését. Ez az egyszerűsítés az olyan feladatnál, melynél az F-beli elemek generálása igen időigényes, különösen gazdaságos: a megoldási idő az átlagos feladatok megoldási

	S5	S7	S10	H3	H6	BR	GP	C6	SCH
Gomulka-Branin	5500	5020	4860						
Törn	3679	3606	3874	2584	2447	1558	2499		
Gomulka-V. M.	7085	6684	7352	6766	11125	1318	1495		
Price	3800	4900	4400	2400	7600	1800	2500		
Fagioli	2504	2509	2518	513	2916	1600	158		
Mockus	1170	1279	1209	513	1232	189	362		
Boender et al.	567	624	755	235	462	235	398		
Boender-Csendes	990	1767	2396	216	1446	330	436	233	
Jones et al.	15	145	145	199	571	195	285	2967	
Storn-Price	6400	6194	6251	476	7220	1190	1018	416	1371
OSCA	3992	3992	3992	1497	1497	1497	1497	1497	1497

2. táblázat. Különböző módszerek összehasonlítása a függvényhívások száma alapján (Dixon & Szegő függvények)

idejére csökken. Nézzünk egy ilyen tesztfeladatot Michalewicz et al. és Binh et al. nyomán ([21], [22], [3]).

$$f(X) = -x_1^2 + x_2^2 + x_1x_2 - 14x_1 - 16x_2 + (x_3 - 10)^2 + 4(x_4 - 5)^2 + (x_5 - 3)^2 + \\ + 2(x_6 - 1)^2 + 5x_7^2 + 7(x_8 - 11)^2 + 2(x_9 - 10)^2 + (x_{10} - 7)^2 + 45$$

feltételek:

$$105 - 4x_1 - 5x_2 + 3x_7 - 9x_8 \geq 0$$

$$-10x_1 + 8x_2 + 17x_7 - 2x_8 \geq 0$$

$$8x_1 - 2x_2 - 5x_9 + 2x_{10} + 12 \geq 0$$

$$3x_1 - 6x_2 - 12(x_9 - 8)^2 + 7x_{10} \geq 0$$

$$-3(x_1 - 2)^2 - 4(x_2 - 3)^2 - 2x_3^2 + 7x_4 + 120 \geq 0$$

$$-x_1^2 - 2(x_2 - 2)^2 + 2x_1x_2 - 14x_5 + 6x_6 \geq 0$$

$$-5x_1^2 - 8x_2 - (x_3 - 6)^2 + 2x_4 + 40 \geq 0$$

$$-0.5(x_1 - 8)^2 - 2(x_2 - 4)^2 - 3x_5^2 + x_6 + 30 \geq 0$$

ahol  $x_i \in [-10, 10]$ ,  $i = 1, 2, \dots, 10$ .

A tesztfeladatot evolúciós technikával, penalty függvény alkalmazásával több változatban megoldották. E nehéz problémánál (amely a „death penalty method”-ot illusztrálja) átlagos megoldásként pl. 25,00-öt kaptak a globális optimum 24,3062 értéke helyett ([3]). OSCA eredménye is hasonló: 24,9. Ezt az eredményt a következő paraméter értékeknél kaptuk:  $h = 0,01$ ,  $itt = 5000$ ,  $delta = 10^{-7}$  és  $it = 15469$ .

Összességében OSCA-t sikeresen teszteltük különböző típusú függvények körében. Folytonos, vagy nem folytonos, szeparálható, ill. nem szeparálható, multimodális, ill. unimodális tulajdonságokkal jellemezhető függvények, ill. feltételrendszerek egyaránt előfordultak a tesztfeladatok körében.

### 3. Néhány polinomiális programozási probléma

Az OSCA, valamint a hozzá tartozó program lehetőségeinek további szemléltetésére válasszuk a polinomiális programozás (PP) témakörét. A PP probléma általános alakja:

$$\text{Global minimuma } \sum_{i=1}^n f_j(X)$$

feltételek

$$g_k(X) \leq 0 \quad k = 1, 2, \dots, q \quad 0 \leq l_i \leq x_i \leq u_i \quad i = 1, 2, \dots, n$$

ahol  $f_j(X)$  és  $g_k(X)$   $X$  polinomiális függvényei és  $l_i$ ,  $u_i$  az alsó és felső korlátját jelöli  $x_i$ -nek.

PP problémára pl. Li és Chang mutat be részletesen 3 feladatot cikkében, mellyel a globális optimum közelítő megoldását egy új linearizáló eljárással szemléltetik [20]. Oldjuk meg e három feladatot az OSCA-val is.

A három feladat:

Globális minimuma

$$PP1(X) = x_1 \cdot x_2 \cdot x_3 + x_1^2 - 2x_1 \cdot x_2 - 3x_1 \cdot x_2 + 5x_2 \cdot x_3 - x_3^2 + 5x_2 + x_3$$

feltételek:

$$4x_1 + 3x_2 + x_3 \leq 20,$$

$$x_1 + 2x_2 + x_3 \geq 1,$$

$$2 \leq x_1 \leq 10,$$

$$0 \leq x_2 \leq 10,$$

$$4 \leq x_3 \leq 8$$

Globális minimuma

$$PP2(X) = 0,6224x_3 \cdot x_4 + 1,7781x_2 \cdot x_3^2 + 3,1661x_1^2 \cdot x_4 + 19,84x_1^2 \cdot x_3$$

feltételek:

$$-x_1 + 0,0193x_3 \leq 0,$$

$$-x_2 + 0,00954x_3 \leq 0,$$

$$-\pi x_3^2 \cdot x_4 - 4/3\pi x_3^3 + 750 \leq 0,$$

$$1,0 \leq x_1 \leq 1,375,$$

$$0,625 \leq x_2 \leq 1,0$$

ahol  $x_1$  és  $x_2$  diszkrét változók 0,0625 lépésközzel, és

$$47,5 \leq x_3 \leq 52,5,$$

$$90,0 \leq x_4 \leq 112,0$$

Globális minimuma

$$PP3(X) = -(x_1^2 + x_2^2 + x_3^2),$$

feltételek:

$$0,5x_1 + x_2 + x_3 - 1 = 0,$$

$$x_1^2 + 2/3x_2^2 + 1/4x_3^2 - 4 = 0$$

ahol  $x_1, x_2, x_3$  tetszőleges értéket felvehet.

A korlátozó feltételekkel kijelölt tartomány az első két feladatnál egyszerű, könnyen lehetővé teszi véletlen mintapontok generálását. A harmadik (PP3) feladatnál az egyenlőségek már olyan tartományt jelölnek ki, amelyben időigényes a véletlen mintapontok generálása. E probléma megoldásánál  $\delta = 10^{-4}$  értéket választottam és feltettem, hogy  $x_1, x_2, x_3 \in [-3, 3]$ .

A PP1 és PP3 feladat esetén Li és Chang linearizáló technikával kapott eredményeivel azonosat, a PP2 feladatnál pedig a korábbi globális optimumnál jobb közelítést kaptam. Az eredményeket, alkalmazott paramétereket a 3. táblázat szemlélteti (csillaggal jelölve a korábbinál jobb eredményt). A PP2 feladat esetén Sangren (1990), Fu et al. (1991), Li et al. (1998) módszereivel és az OSCA-val kapott eredmények összehasonlítását a 4. táblázat szemlélteti.

	Dim	$h$	$itt$	$delta$	$it$	eredmény
PP1	3	0.01	2000	$10^{-3}$	3493	-119
PP2	4	0.01	2500	$10^{-5}$	8982	6395,5*
PP3	3	0.01	500	$10^{-5}$	11477	-10,99

3. táblázat. A PP problémák megoldása  
( $eps = 0,001$ ,  $kn = 500$ ,  $t = 2$ )

változó/módszer	Sangren's	Fu et al.	Li et al.	OSCA
$x_1$	1,125	1,125	1,000	1,000
$x_2$	0,625	0,625	0,625	0,625
$x_3$	48,95	48,35	50,25	47,5
$x_4$	106,72	11,745	90,991	90,00
PP2 (X)	7982,5	8048,6	7127,3	6395,5

4. táblázat. PP2 megoldásának összehasonlítása más módszerekkel

A teszt feladatokkal kapcsolatban ki kell emelnem, hogy feltételrendszerekkel adott problémák esetén a rövid megoldási idők egyértelműen a  $D(a_i)$  távolság alkalmazásának köszönhetők. Ennek szemléltetésére vizsgáljunk meg egy másik megoldást. Módosítsuk a mintapontok generálását és csak olyan véletlen pontokat engedjünk figyelembe venni, amelyek a korlátozó feltételnek megfelelnek (azaz  $a_i \in F$ ). Ez a változat, amennyiben  $F$  pontjait nem nehéz megtalálni, az előzőhöz hasonló idő alatt, ugyanazokat a jó megoldásokat nyújtja. Ha viszont  $F$  pontjait nagyon nehéz megtalálni, a megoldás időigénye nagyon megnőhet.

A  $D(a_i)$  értékek alapján való vizsgálat időigénye egységesen rövid: pl. a PP3 feladatnál 15 percről 2 percre, Michalewicz teszt feladatánál pedig több óráról 3 percre csökkent a megoldás ideje egy 300 Mhz-es Pentium II-es PC-n.

#### 4. Összefoglalás

E cikkben egy klaszterező algoritmust mutattunk be multimodális függvények optimalizálására. Az OSCA eltérően a szokásos módszerektől, folyamatosan finomítva alakítja ki a klasztereket az egyes lokális minimumok köré. A feltételrendszereket az evolúciós algoritmusoknál alkalmazott mérték felhasználásával a célfüggvénytől elkülönítve kezeli. A benchmark-ként alkalmazott függvények szélsőértékeit általában 0,01 pontossággal határozza meg, és a konvergencia, a függvény kiértékelések száma hasonló más klaszterező algoritmusokéval. Az OSCA a szokásos sztochasztikus módszereknél szélesebb körben alkalmazható: több lokális minimum (pl. 20–30) keresésére, magasabb, 20–30 dimenziós feladatok megoldására, ill. feltételrendszerek kezelésére is gazdaságosan használható.

## Irodalom

- [1] Agapie, A. and Caragead, D., Genetic Algorithms, Schemata Construction and Statistics, in: *Computational Intelligence*, ed. B. Reusch (Springer Verlag, Berlin, 1997) 16–23.
- [2] Bazaraa, M. S. and Shetty, C. M., *Nonlinear programming*, John Wiley & Sons (New York, 1979)
- [3] Binh, T. T. and Korn, U., Scalar Optimization with Linear and Nonlinear Constraints Using Evolution Strategies, in: *Computational Intelligence*, ed. B. Reusch Springer Verlag (Berlin, 1997), 381–392.
- [4] Björkman, M. and Holmström K., „Global Optimization Using the DIRECT Algorithm in MATLAB”, Technical Report IMg-TOM-1998-12. (Sweden, 1998).
- [5] Boender, C. G. E., Rinnooy Kan, A. H. G., Strongie, L. and Timmer, G. I., A stochastic method for global optimization, *Math. Progr.*, **22** (1982), 125–140.
- [6] Boender, C. G. E. and Romeijn, H. E., Stochastic methods, in: *Handbook of Global Optimization*, Horst R., Pardalos, (eds.) Kluwer (Dordrecht, 1995), pp. 829–869.
- [7] Borgulya, I., Két fuzzy osztályozó módszer, *Sigma*, **XXIX**. No. 1–2. (1998), 7–28.
- [8] Borgulya, I., Multimodal Function Optimization with a Clustering Algorithm (1999), Manuscript.
- [9] Borgulya, I., Constraint Optimization with a Clustering Algorithm, *Central European Journal for Operation Research*, in print.
- [10] Csendes, T., Nonlinear parameter estimation by global optimization — efficiency and reliability, *Acta Cybernetica*, **8** (1988), 361–370.
- [11] Csendes, T., Egy klaszterező globális optimalizálási módszer a paraméterbecslési feladat megoldására, *Sigma*, **XXIII**. No. 1–2. (1992), 17–35.
- [12] Dixon, L. C. W. and Szegő, G. P. (eds), *Towards Global Optimization 2*, North-Holland (Amsterdam, 1978).
- [13] Fu, J. F., et al., A mixed integer-discrete-continuous programming method and its application to engineering design optimization, *Engineering Optimization*, **17** (1991), 263–280.
- [14] Hansen, E. R., *Global Optimization Using Interval Analysis*, Dekker (New York, 1992).
- [15] Hocaoglu, C. and Sanderson, A., Evolutionary Speciation Using Minimal Representation Clustering, in: *Evolutionary Programming*, McDonnell Z. R., Reynolds, R. G., Fogel D. B. (eds.) Proc. of the 4<sup>th</sup> Annual Conference on Evolutionary Programming (The MIT Press Cambridge, 1995), pp. 187–203.
- [16] Horst, R. and Tuy, H., *Global Optimization: Deterministic Approaches*, Springer-Verlag (Berlin, 1993).
- [17] Huyer, W. and Neumaier, A., *Global optimization by multilevel coordinate search*, (Report (FWF) P11516-MAT, University Wien, 1999).
- [18] Ingber, A. L., Adaptive simulated annealing (ASA): Lessons learned, *Control Cybernet*, **25** (1996), 33–55.
- [19] Jones, D. R., Perttunden, C. D. and Stuckman, B. E. Lipschitzian optimization without the Lipschitz Constant, *J. of Optimization Theory and Application*, **79** (1993), 157–181.
- [20] Li, H., Chang, C., An approximate approach of global optimization for polynomial programming problems, *European Journal of Operational Research*, **107** (1998), 625–632.
- [21] Michalewicz, Z. Heuristic methods for evolutionary computation techniques, *Journal of Heuristics*, **1**(2) (1995), pp. 177–206.

- [22] Michalewicz, Z., Dasgupta, D., Riche, R. and Schoenauer, M., Evolutionary algorithms for constrained engineering problems, in: *Proc. of the 6<sup>th</sup> Conf. on Parallel Problems Solving from Nature*, Davidor V., Schwefel H. P., Manner R. (Eds.), Springer Verlag (Berlin, 1996).
- [23] Nissen, V. *Einführung in Evolutionäre Algorithmen*, Vieweg (Braunschweig, 1998).
- [24] Pérowski, A., A New Selection Operator Dedicated to Speciation, in: *Proc. of the 7th International Conf. on Genetic Algorithms*, Back (ed.), Morgan Kaufman Pub. Inc. (San Francisco, 1997), pp. 144–151.
- [25] Pintér, J. D. *Global Optimization in Action*, Kluwer (Dordrecht, 1995).
- [26] Richardson, J., Palmer, M., Liepins, G. and Hilliard, M. Some guidelines for genetic algorithms with penalty functions, *Proc. of the 3<sup>rd</sup> Int. Conf. on Genetic Algorithms*, Los Altos, CA (Morgan Kaufmann Pub., 1989), pp. 19–197.
- [27] Sangren, E., Nonlinear integer and discrete programming in mechanical design optimization, *Journal of Mechanical Design*, **112** (1990), 223–229.
- [28] Solis, F. J. and Wets, R. J. B., Minimization by random search techniques, *Mathematics of Operations Research*, **6** (1981), 19–30.
- [29] Storn, R. and Price, K., *Differential evolution — a simple and efficient adaptive scheme for global optimization over continuous spaces* (Technical Report TR-95-912, ICSI, 1995).
- [30] Törn, A. A search clustering approach to global optimization, in: *Towards Global Optimization 2*, Dixon L. C. W., Szegő G. P. (eds.), North-Holland (Amsterdam, 1978).
- [31] Törn, A. A. and Zilinskas, A. *Global Optimization*. Springer Verlag (Berlin, 1989).
- [32] Törn, A. A. and Viitanen, S. Topographical global optimization using presampled points, *Global Optimization*, **5** (1994), 267–276.
- [33] Zangwill, W. I., The Convex Simplex Method, *Management Science*, **13** (1967), pp. 221–283.
- [34] Zoutendijk, G., *Methods of Feasible Directions*, Elsevier (Amsterdam and New York, 1960).

Beérkezett: 1999. november 22.

BORGULYA ISTVÁN  
PÉCSI TUDOMÁNYEGYETEM  
KÖZGAZDASÁGTUDOMÁNYI KAR  
GAZDASÁGI INFORMATIKA TANSZÉK  
H-7621 PÉCS  
RÁKÓCZI ÚT 80.  
E-mail: BORGULYA@GALILEO.KTK.PTE.HU

SOME POLYNOMIAL PROGRAMMING PROBLEMS SOLVING WITH  
A CLUSTERING ALGORITHM

ISTVÁN BORGULYA

In my study I describe an optimum search clustering algorithm and I demonstrate its capabilities by the approximate solving of a few polynomial programming problems. This clustering algorithm determines in a given number the local or global minima of the multimodal functions as the center of the clusters. Unlike normal clustering algorithms it approaches the extrema continuously in a single phase, on the basis of the sample points. The method determines the values and places of the extrema of the usual test examples with an accuracy of about 0.01. The OSCA can be used to solve linear or nonlinear constrained problems, and as the examples demonstrate it gives a good approximation of the global optimum in case of benchmark optimization problems as well.



# DEKOMPOZÁBILIS HALMAZOK\*

KÁNNAI ZOLTÁN

Budapest

A dekompozabilitás mint a konvexitás egy analógiájára épülő elmélet mára már alapvető szereppel bír mind az optimális irányítások, mind a differenciáltartalmazások elméletében. Legyen  $T$  mértéktér,  $E$  Banach-tér és  $L^1(T, E)$  jelölje az  $E$  fázisterű, a  $T$  téren Lebesgue-integrálható függvények Banach-terét. Egy  $M \subseteq L^1(T, E)$  halmazt *dekompozábilis*nak nevezzük, ha minden  $f, g \in M$  függvényre és  $A \subseteq T$  mérhető halmazra

$$\chi_A \cdot f + (1 - \chi_A) \cdot g \in M.$$

Dekompozábilis halmazra tipikus példa egy mérhető halmazértékű leképezés mérhető szelekcióinak összessége. Ugyanakkor a konvex halmazokra közismert állítások közül egy sereg dolog átvihető dekompozábilis halmazokra, pl. fixponttételek, folytonos szelekciók; amelyek nemcsak önmagukban érdekesek, hanem hatékonyan alkalmazhatók a kezdetiérték-feladatok perturbációiban, illetve a nemlineáris irányítások elméletében. Sőt segítségükkel kézenfekvő módon adódik a Ljapunov-tétel és a bang-bang elv egy közös direkt általánosítása.

## 1. Bevezetés

A dekompozabilitás fogalma a matematikai szakirodalomban alig több, mint húsz éve jelent meg [4], és bár a dekompozabilitás és konvexitás közötti párhuzam nagyon kézenfekvő, a dekompozábilis halmazokra vonatkozó súlyosabb matematikai eszközök megjelenése alig tíz éves múltra tekint vissza. Ezen dolgozatban a dekompozabilitást úgy mutatjuk be mint az analízisben alapvető szerepet betöltő koncepciót. Ebben, továbbá a második, harmadik és negyedik paragrafusban a dekompozábilis halmazoknak a matematikai analízis különböző területein vett alkalmazásaiból már korábban publikált eredményeket vonultatunk fel, míg az utolsó paragrafusban az Aumann-integrálok és optimális irányítások elméletéből korábban ismert eredmények új, absztraktabb változatait bizonyítjuk a dekompozábilis halmazok által kínált elegáns megközelítésben.

---

\*Ez a dolgozat a XVII. Magyar Operációkutatási Konferencián (Veszprém, 1999. október 7–9.) elhangzott előadás írott változata.

A konvexitás és dekompozáltság elméletének analógiája a konvex és dekompozáltság kombináció analógiáján alapul, melyre az egész elmélet ráépül:

**Konvex kombináció:**

$$x, y \in X, \lambda \in [0, 1] \longrightarrow \lambda \cdot x + (1 - \lambda) \cdot y$$

**Dekompozáltság kombináció:**

$$f, g \in L^1(T, E), A \subseteq T \text{ mérhető} \longrightarrow \chi_A \cdot f + (1 - \chi_A) \cdot g$$

ahol persze  $X$  valós vektortér,  $E$  Banach-tér és  $T$  mértéktér. Mind a konvex, mind a dekompozáltság halmazokra vonatkozó legjelentősebb analízisbeli eszközök szoros kapcsolatban állnak a halmazértékű analízissel, folytonossági és mérhetőségi szempontból egyaránt. A következőkben szükségünk lesz az ezekhez kötődő legalapvetőbb fogalmakra.

Legyenek  $X$  és  $Y$  topologikus terek, jelölje  $\mathcal{P}(Y)$  az  $Y$  hatványhalmazát. Egy  $F : X \rightarrow \mathcal{P}(Y)$  leképezést

- *alsó-Vietoris-folytonosnak* (a-V-folytonos) mondunk, ha minden  $G \subseteq Y$  nyílt halmazra

$$\{x \in X : F(x) \cap G \neq \emptyset\}$$

nyílt halmaz  $X$ -ben;

- *felső-Vietoris-folytonosnak* (f-V-folytonos) mondunk, ha minden  $G \subseteq Y$  nyílt halmazra

$$\{x \in X : F(x) \subseteq G\}$$

nyílt halmaz  $X$ -ben.

- Tegyük fel, hogy  $X$  nem topologikus, hanem mérhető tér. Az  $F$  leképezést *mérhetőnek* nevezzük, ha minden  $G \subseteq Y$  nyílt halmazra

$$\{x \in X : F(x) \cap G \neq \emptyset\}$$

mérhető halmaz  $X$ -ben.

Legyen  $T$  véges atommentes mértéktér,  $E$  Banach-tér és  $L^1(T, E)$  jelölje az  $E$  fázisterű, a  $T$  téren Lebesgue-integrálható függvények Banach-terét. Egy  $M \subseteq L^1(T, E)$  halmazt *dekompozáltságnak* nevezzük, ha minden  $f, g \in M$  függvényre és  $A \subseteq T$  mérhető halmazra

$$\chi_A \cdot f + (1 - \chi_A) \cdot g \in M.$$

**Konvexitási eredmények:**

Legyen  $X$  metrikus tér,  $E$  Banach-tér és  $K \subseteq E$  konvex.

1. Schauder-fixponttétel: Legyen  $K$  zárt, ekkor tetszőleges  $f : K \rightarrow K$  folytonos relatív kompakt értékkészletű függvénynek van fixpontja.
2. Dugundji-féle kiterjesztési tétel: Bármely  $M \subseteq X$  zárt halmazon értelmezett tetszőleges  $f : M \rightarrow K$  folytonos függvény folytonosan kiterjeszthető  $X \rightarrow K$  folytonos függvényre.
3. Michael-féle szelekciós tétel: Minden  $F : X \rightarrow \mathcal{P}(E)$  nemüres konvex zárt értékű a-V-folytonos leképezésnek van folytonos szelekciója.
4. Cellina-féle szelekciós tétel: Tetszőleges  $F : X \rightarrow \mathcal{P}(E)$  nemüres konvex zárt értékű f-V-folytonos leképezésnek  $\forall \varepsilon > 0$  esetén van folytonos  $\varepsilon$ -gráfszelekciója, azaz  $f : X \rightarrow E$  folytonos függvény, amelyre

$$\text{graph } f \subseteq B(\text{graph } F, \varepsilon).$$

**Dekompozabilitási eredmények:**

A fentiekben az  $E$  teret az  $L^1(T, E)$  térrel, a *konvex* szót a *dekompozábilis* szóval helyettesítve mind a négy állítás igaz marad (A. Bressan–G. Colombo [2]).

**2. Dekompozabilitás és szélsőérték**

Szélsőértékfeladatok gyakran a következő típusú feladatra vezetnek: egy normált függvénytérben egy zárt halmazon kívüli ponthoz találjunk (ha van) a zárt halmazban legközelebb fekvő pontot. A hétköznapi gondolkodás azt sugallná, hogy egy halmazhoz legközelebbi függvényt pontonként (változónként) próbáljunk megtalálni. Ám néhány pontban optimális függvény az eredeti feladat szempontjából esetleg nagyon is rossz lehet. Ami pedig esetleg jó, az pontonként nem bizonyul a legjobbnak. Dekompozábilis halmazok esetén viszont a két szemléletű optimum lényegében egybeesik:

1. TÉTEL. Legyen  $M \subseteq L^1(T, E)$  nemüres dekompozábilis halmaz,  $u \in M$  és  $v \in L^1(T, E)$ .  $v$ -nek az  $u$ -tól vett  $L^1$ -beli távolsága pontosan akkor minimális, ha minden  $w \in M$ -re

$$\|u(t) - v(t)\| \leq \|w(t) - v(t)\| \quad \text{m.m. } t \in T \text{ esetén.}$$

*Bizonyítás.* A  $\|w(\cdot) - v(\cdot)\|$  ( $w \in M$ ) függvények dekompozábilis halmazt alkotnak  $L^1(T, \mathbb{R})$ -ben. Ha ennek  $\|u(\cdot) - v(\cdot)\|$  majdnem mindenütt minimális eleme, akkor persze minimális normájú is. Most azt tegyük föl, hogy  $\|u(\cdot) - v(\cdot)\|$  minimális normájú. Ha majdnem mindenütt nem lenne minimális, akkor volna olyan

$w \in M$ , hogy egy pozitív mértékű  $A$  halmazon  $\|u(t) - v(t)\| > \|w(t) - v(t)\|$  állna fenn. Ekkor  $z = \chi_A \cdot w + (1 - \chi_A) \cdot u \in M$  és  $\|u - v\|_{L^1} > \|z - v\|_{L^1}$ , ami ellentmondás.  $\square$

A dekompozálható halmazoknak a mérhető szelekciók elméletében is fontos szerep jut, ami pedig az optimális irányítások elméleti vizsgálataiban bír alapvető szereppel. 1977-ben F. Hiai és H. Umegaki [4] bizonyították, hogy szeparálható  $L^1(T)$  és  $E$  esetén egy  $M \subseteq L^1(T, E)$  halmaz pontosan akkor dekompozálható és zárt, ha van olyan  $F : T \rightarrow \mathcal{P}(E)$  zárt értékű mérhető leképezés, hogy

$$M = \{u \in L^1(T, E) : u(t) \in F(t) \text{ m.m.t } t \in T \text{ esetén}\}.$$

Az idén Páles Zsolt és Vera Zeidan bizonyították hasonló állítást az  $L^\infty$  térben [7].

### 3. Differenciáltartalmazások

A differenciáltartalmazások egzisztenciátételeiben nagyon szemléletesen mutatkozik meg a konvexitás és dekompozálhatóság közötti párhuzam. Differenciáltartalmazásnak az

$$x'(t) \in F(t, x(t)) \quad \text{m.m.t esetén}$$

alakú feladatokat szokás nevezni (a keresett függvény abszolút folytonos valamely intervallumon). A differenciáltartalmazásoknak két fő általános típusa van:

- a jobboldal konvex zárt értékű és f-V-folytonos;
- a jobboldal (nem feltétlenül konvex) zárt értékű és a-V-folytonos.

Az előbbiek esetében a fentebb említett konvexitási eredmények jelentik a fő eszköztárat, míg az utóbbiak esetében ugyanezek dekompozálható analogonjai. Valóban,

$$\Gamma : C([a, b], E) \rightarrow L^1([a, b], E)$$

$$\Gamma(f) := \{u \in L^1([a, b], E) : u(t) \in F(t, f(t)) \text{ m.m.t } t \in [a, b] \text{ esetén}\}$$

az első esetben konvex értékű, a másodikban nem feltétlenül, de dekompozálható értékű. Ha ezt a párhuzamot tovább vesszük, akkor arra jutunk, hogy irányításelméletben lineáris és nemlineáris irányítási feladatok úgy felelnek meg egymásnak, mint absztrakt esetben a konvex és dekompozálható halmazok.

A fenti transzformációt alkalmazva V. V. Goncharov [3] bizonyított először nemkonvex jobboldalú differenciáltartalmazásra egzisztenciátételt.

#### 4. Kontrakció és szelekció

A dekompozábilis értékű leképezések szelekciós problematikája tovább szőhető, és egybe szőhető a fixpont problematikájával. Ez differenciáltartalmazások perturbációelméletében jelent alkalmazást. Legyen

$$F : X \times L^1([a, b]; E) \rightarrow \mathcal{P}(L^1([a, b]; E))$$

zárt dekompozábilis értékű leképezés, melyre

(i) van olyan folytonos  $\alpha : X \rightarrow \mathbb{R}_+$  függvény, hogy

$$F(x, u) \subseteq \alpha(x) \cdot \mathbf{B}_{L^1([a, b]; E)}$$

minden  $x \in X$  és  $u \in L^1([a, b]; E)$  esetén;

(ii) minden  $y \in L^1([a, b]; E)$  esetén

$$F(., y) : X \rightarrow \mathcal{P}(L^1([a, b]; E))$$

a-V-folytonos leképezés;

(iii) Van olyan korlátos és folytonos  $\gamma : X \rightarrow L^1([a, b]; \mathbb{R})$  függvény, hogy minden  $x \in X, y_1, y_2 \in L^1([a, b]; E)$  és  $u \in F(x, y_1)$  esetén  $\exists v \in F(x, y_2)$  melyre

$$\|u(t) - v(t)\| \leq \gamma(x)(t) \cdot \int_a^t \|y_1(\tau) - y_2(\tau)\| d\tau$$

m.m.t  $\in [a, b]$ -re.

2. TÉTEL. (i)–(iii) teljesülése esetén van olyan

$$f_0 : X \rightarrow L^1([a, b]; E)$$

folytonos függvény, melyre

$$f_0(x) \in F(x, f_0(x))$$

minden  $x \in X$  esetén.

A bizonyítás [6]-ban található.

Legyen most  $S$  szeparábilis metrikus tér és  $E$  valós Banach-tér. Tekintsünk egy zárt értékű

$$F : [0, 1] \times E \times S \rightarrow \mathcal{P}(E)$$

leképezést az alábbi tulajdonságokkal:

(a) minden  $(x, \xi) \in E \times S$  mellett  $F(., x, \xi)$  mérhető;

(b) van olyan folytonos  $\alpha : S \rightarrow L^1([0, 1]; \mathbb{R})$  függvény, hogy minden  $(x, \xi) \in E \times S$  és m.m.t  $\in [0, 1]$  mellett

$$F(t, x, \xi) \subseteq \alpha(\xi)(t) \cdot B_E;$$

(c) van olyan folytonos  $\gamma : S \rightarrow L^1([0, 1]; \mathbb{R})$  függvény, hogy m.m.t  $\in [0, 1]$ , minden  $x, y \in E$  és minden  $\xi \in S$  esetén

$$\mathcal{D}_{\text{Hausdorff}}(F(t, x, \xi), F(t, y, \xi)) \leq \gamma(\xi)(t) \cdot \|x - y\|;$$

(d) m.m.t  $\in [0, 1]$  és minden  $x \in E$  mellett

$$F(t, x, \cdot) : S \rightarrow \mathcal{F}(E)$$

a-V-folytonos leképezés.

3. KÖVETKEZMÉNY. Tegyük föl az (a)–(d) feltételek teljesülését. Ekkor van olyan folytonos  $x : E \times S \rightarrow AC([0, 1]; E)$  függvény, hogy minden  $(z, \xi) \in E \times S$  mellett  $x(z, \xi) \in AC([0, 1]; E)$  megoldása a

$$(1) \quad \begin{cases} (x(z, \xi))'(t) \in F(t, x(z, \xi)(t), \xi) \text{ m.m.t} \\ x(z, \xi)(0) = z \end{cases}$$

Cauchy-feladatnak.

A következménynek a tételre való visszavezetése a

$$\Gamma((z, \xi), f) := \left\{ u \in L^1([0, 1]; E) : u(t) \in F\left(t, z + \int_0^t f, \xi\right) \text{ m.m.t } \in [0, 1] \right\}.$$

leképezés bevezetésével történik (lásd szintén [6]).

## 5. Függvényhalmazok integrálása

A továbbiakban legyen  $(T, \mathcal{S}, \mu)$  atommentes mértéktér,  $E$  Banach-tér,  $H$  véges dimenziós euklideszi tér és  $\Phi : L^1(T, E) \rightarrow H$  folytonos lineáris operátor. Jelölje  $M \subseteq L^1(T, E)$  esetén

$$\int_T M d\mu := \left\{ \int_T u d\mu : u \in M \right\}.$$

Az  $\int_T M d\mu$  halmazt az  $M$  halmaz integráljának mondjuk. Megjegyezzük, hogy egy  $F : T \rightarrow \mathcal{P}(E)$  leképezésre

$$M_F := \{u \in L^1(T, E) : u(t) \in F(t) \text{ m.m. } t \in T\}$$

definícióval  $\int_T M_F d\mu$  éppen  $F$  Aumann-integrálja.

Egy  $M \subseteq L^1(T, E)$  halmazt erősen konvexnek nevezünk, ha minden  $u, v \in M$  és  $h : T \rightarrow [0, 1]$  mérhető függvényre  $h \cdot u + (1 - h) \cdot v \in M$ . Egy erősen konvex halmaz egyszersmind konvex és dekompozábilis. Egy  $N \subseteq L^1(T, E)$  halmaz erős konvex burkát jelölje  $\text{sco}(N)$ .

4. TÉTEL. *Tetszőleges  $M \subseteq L^1(T, E)$  dekompozábilis halmazra*

$$\Phi(M) = \Phi(\text{sco}(M)),$$

*speciálisan  $\Phi(M)$  konvex halmaz.*

*Bizonyítás.* Jelölje  $N \subseteq L^1(T, E)$  esetén

$$c(N) := \{h \cdot u + (1 - h) \cdot v : u, v \in N \text{ és } h : T \rightarrow [0, 1] \text{ mérhető}\}.$$

Ha  $N$  dekompozábilis, akkor  $c(N)$  is triviálisan dekompozábilis. Először megmutatjuk, hogy bármely  $N$  dekompozábilis halmazra

$$\Phi(N) = \Phi(c(N)).$$

A baloldal nyilván része a jobboldalnak. Legyen  $y \in \Phi(c(N))$ . Ekkor vannak olyan  $u, v \in N$  és  $h : T \rightarrow [0, 1]$  mérhető függvények, hogy  $y = \Phi(h \cdot u + (1 - h) \cdot v)$ . A

$$\nu(A) = \Phi(\chi_A \cdot (u - v))$$

módon értelmezett  $\nu : \mathcal{S} \rightarrow H$  leképezés a Lebesgue-féle konvergenciatételből adódólag mérték, melynek  $\mathcal{R}$  értékkészlete Ljaponov tétele értelmében konvex és kompakt halmaz.  $\nu$  nyilván abszolút folytonos a véges és atommentes

$$\lambda(A) = \int_A \|u - v\|_E d\mu$$

mértékre nézve. Jelölje  $g$  a  $\nu$  mérték  $\lambda$ -ra vonatkozó Radon–Nikodým-deriváltját. Indirekt tegyük föl, hogy  $\Phi(h \cdot (u - v)) \notin \mathcal{R}$ . Ekkor van olyan  $p \in H^*$ , hogy

$$\int_T phg d\lambda = p \int_T hg d\lambda = p\Phi(h(u - v)) > \max_{x \in \mathcal{R}} px.$$

No de  $p \cdot \nu$  értékkészlete is konvex kompakt, azaz kompakt intervallum, amelynek  $\int_A pg \, d\lambda$  és  $\int_B pg \, d\lambda$  elemei, ahol  $A$  illetve  $B$  a  $pg$  függvény pozitivitási illetve negatívítási nívóhalmaza. Másrészt nyilván  $\int_T phg \, d\lambda$  ezen két érték közé esik, tehát eleme  $p \cdot \nu$  értékkészletének, ami ellentmondásban áll a fenti egyenlőtlenséggel. Tehát  $\Phi(h \cdot (u - v)) \in \mathcal{R}$ . Ezért van olyan  $C \in \mathcal{S}$ , hogy  $\Phi(h \cdot (u - v)) = \Phi(\chi_C \cdot (u - v))$ . Innen

$$y = \Phi(h \cdot u + (1 - h) \cdot v) = \Phi(\chi_C \cdot u + (1 - \chi_C) \cdot v) \in \Phi(N).$$

Tehát  $\Phi(N) = \Phi(c(N))$ . Defináljuk most rekurzióval a  $c^n(M)$  halmazokat ( $n \in \mathbb{N}$ ) a következő módon:  $c^1(M) = c(M)$ ,  $c^{n+1}(M) = c(c^n(M))$ . A fentiek értelmében mindegyik  $c^n(M)$  dekompozálható és teljes indukcióval adódik, hogy

$$\Phi(c^n(M)) \, d\mu = \Phi(M).$$

Innen

$$\Phi(\text{sco}(M)) = \Phi\left(\bigcup_{n \in \mathbb{N}} c^n(M)\right) = \Phi(M).$$

□

5. KÖVETKEZMÉNY. Tetszőleges  $M \subseteq L^1(T, H)$  dekompozálható halmazra  $\int_T M \, d\mu$  konvex halmaz.

6. MEGJEGYZÉS. Legyen  $F : T \rightarrow \mathcal{P}(E)$  kompakt értékű mérhető és integrálhatóan korlátos leképezés. A Kuratowski–Ryll–Nardzewski-féle szelekciós tétel [1] alapján könnyen látható, hogy

$$M_{\text{conv}} F = \text{sco}(M_F).$$

Mindkét oldalt integrálva és a fenti tételt az integrálásra alkalmazva nyerjük Richter tételét ([5] 8.2.):

$$\int_T F \, d\mu = \int_T \text{conv } F \, d\mu.$$

Végül egy olyan tételt közlünk, melynek triviális alkalmazásával igazolható a lineáris bang-bang elv (az alkalmazást az olvasóra bizzuk). Egy  $M \subseteq L^1(T, E)$  konvex halmaz extrémis pontjainak halmazát jelölje  $\text{ext}(M)$ .

7. TÉTEL (absztrakt bang-bang elv). Legyen  $M \subseteq L^1(T, E)$  erősen konvex és gyengén kompakt halmaz. Ekkor

$$\Phi(M) = \Phi(\text{ext}(M)).$$



*Bizonyítás.* A jobboldal triviálisan része a baloldálnak. Legyen most  $x \in \text{ext}(\Phi(M))$ . A Krein–Milman-tétel alapján a nemüres konvex és gyengén kompakt

$$\Phi^{-1}(x) \cap M$$

halmaznak van egy  $u$  extrémális pontja. Persze ekkor  $u \in \text{ext}(M)$ , innen  $x \in \Phi(\text{ext}(M))$ . Tehát

$$\text{ext}(\Phi(M)) \subseteq \Phi(\text{ext}(M)).$$

$\text{ext}(M)$  triviálisan dekompozábilis halmaz. Tehát a 4. Tétel miatt  $\Phi(\text{ext}(M))$  konvex. Ezért a Krein–Milman-tétel véges dimenziós változata miatt

$$\Phi(M) = \text{conv}(\text{ext}(\Phi(M))) \subseteq \Phi(\text{ext}(M)). \quad \square$$

8. MEGJEGYZÉS. Ha  $\mu : S \rightarrow H$  mérték, akkor

$$M = \left\{ f \in L^1(T, |\mu|) : \mathcal{R}_f \subseteq [0, 1] \right\} \quad \text{és} \quad \Phi(f) = \int_T f d\mu$$

választással a fentiekből triviálisan visszkapjuk Ljapunov tételét. Tehát az absztrakt bang-bang elv a Ljapunov-tétel direkt általánosításának tekinthető.

9. MEGJEGYZÉS. Ha a fentiekbeli  $H$  nem véges dimenziós, hanem  $E$ -vel egyezik meg (ami tetszőleges Banach-tér), akkor  $\Phi$  gyanánt az integrálást választva még mindig igaz marad a következő: bármely  $u \in L^1(T, E)$  függvény  $L^1$ -ben közelíthető lépcsős függvényekkel. Ezek integrálmértékeinek konvex kompakt értékkészletei Hausdorff-metrikában tartanak az  $u$  integrálmértékének értékkészletéhez. Emiatt  $\int_T u d\mu$  még mindig konvex és kompakt. Innen a fentieket végigjárva továbbra is igaz lesz, hogy egy  $M \subseteq L^1(T, E)$  dekompozábilis halmazra  $\overline{\int_T M d\mu} = \overline{\int_T \text{sco}(M) d\mu}$ , ahonnan tetszőleges  $M \subseteq L^1(T, E)$  erősen konvex és gyengén kompakt halmazra

$$\int_T M d\mu = \overline{\int_T \text{ext}(M) d\mu},$$

azaz  $\int_T \text{ext}(M) d\mu$  sűrű részhalmaza  $\int_T M d\mu$ -nek. Ez azt jelenti, hogy annyi végtelen dimenzióban is igaz marad, miszerint ha abban az értelemben nem is helyettesíthetünk egy irányítást extrémális irányítással, hogy ugyanabba a pontba jussunk el, de abban az értelemben igen, hogy a kitűzött pontot tetszőlegesen megközelíthetjük.

## Hivatkozások

- [1] Aubin, J.-P. and Cellina, A., *Differential Inclusions*, Springer-Verlag (Berlin, 1984).
- [2] Bressan, A. and Colombo, G., Extensions and selections of maps with decomposable values, *Studia Math.*, **90** (1988), 69–86.
- [3] Goncharov, V. V., On existence of solutions of some type of differential inclusions on a compact set, *Siberian Math. Journal*, **31**(5) (1990), 24–30.
- [4] Hiai, F. and Umegaki, H., Integrals, conditional expectations, and martingales of multivalued functions, *J. Multivariate Analysis*, **7** (1977), 149–182.
- [5] Ioffe, A. D. and Tichomirov, V. M., *Theory of Extremal Problems*, Nauka (Moscow, 1974) (in Russian).
- [6] Kánnai, Z., Contraction-selection theorem and perturbed Lipschitz inclusions, *Pure Math. and Appl.* Vol. 4 (1993) No.4, 479–491.
- [7] Páles, Zs. and Zeidan, V., Characterization of  $L^1$ -closed decomposable sets in  $L^\infty$ , *J. Math. Anal. Appl.* (to appear).

Beérkezett: 1999. november 20.

KÁNNAI ZOLTÁN  
 BUDAPESTI KÖZGAZDASÁGTUDOMÁNYI ÉS  
 ÁLLAMIGAZGATÁSI EGYETEM  
 MATEMATIKA TANSZÉK  
 1092 BUDAPEST, KINIZSI U. 1–7.  
 E-mail: KANNAI@MATH.BKE.HU

## DECOMPOSABLE SETS

ZOLTÁN KÁNNAI

Decomposability, a theory analogous to convexity, has recently achieved a major importance in optimal control and differential inclusions. A set  $M$  in the space of integrable functions  $L^1(T)$  is said to be *decomposable* if for all  $f, g \in M$ ,

$$\chi_A \cdot f + (1 - \chi_A) \cdot g \in M$$

for every measurable subset  $A$  of  $T$ . Several results on convexity can be transferred to decomposability (e.g. existence of continuous selections, fixed points, and continuous perturbations), and they lead to nontrivial and useful results in optimization theory. Moreover, by integrating decomposable sets in  $L^1$ , we get an abstract formulation of the linear bang-bang principle, which is a certain generalization of Lyapunov's range theorem, too.

## MULTIFÁKKAL ADOTT VALÓSZÍNŰSÉGI KORLÁTOK

BUKSZÁR JÓZSEF

Miskolc

Felső korlátokat adunk meg véges számú esemény uniójának valószínűségére néhány, legfeljebb  $m + 1$  eseményt tartalmazó metszet valószínűségei felhasználásával. A korlátokat speciális hipergráfok, az  $m$ -multifák segítségével fogjuk számítani. A dolgozat fő tétele D. Hunter, K. J. Worsley, Prékopa A. és a szerző korábbi eredményeinek közös általánosítása. Az új korlátok hatékonyságát többdimenziós normális eloszlásfüggvények értékeire adott alsó becslések tesztjein mutatjuk be.

### 1. Bevezetés

Legyenek  $A_1, \dots, A_n$  tetszőleges események. Célunk, hogy adott  $m$  pozitív egész mellett felső korlátokat adjunk a  $P(A_1 \cup \dots \cup A_n)$  valószínűségekre a  $P(A_{k_1} \cap \dots \cap A_{k_i})$  ( $1 \leq k_1 < \dots < k_i \leq n$ ,  $i = 1, \dots, m + 1$ ) valószínűségeket – illetve ezek némelyikét – felhasználva. Ezt a Boole [1] által felvetett problémát számos kutató vizsgálta. Részletes történeti áttekintés a [19]-ben található.

Adott  $m$  mellett bevezetjük bizonyos speciális hipergráfok, az  $m$ -multifák fogalmát. Minden egyes  $n$ -csúcú  $m$ -multifához hozzárendelünk egy felső korlátot. Az  $m = 1$  esetben a Hunter–Worsley korlátot (lásd [6] ill. [22]),  $m = 2$  esetben pedig a cseresznyefa korlátot (lásd [3]) kapjuk vissza. Látni fogjuk, hogy tetszőleges  $m$ -multifához meg tudunk adni olyan  $(m + 1)$ -multifát, mely segítségével számított korlát legalább olyan jó és általában jobb, mint az  $m$ -multifa segítségével számított. Tehát az  $m$  növelésével a Hunter–Worsley korláton illetve a cseresznyefa korláton javítani tudunk azon az áron, hogy újabb metszetvalószínűségeket használunk fel.

Bizonyítjuk, hogy egy  $m$ -multifát meghatároz a csúcsainak és éleinek halmaza. Tehát a korlátok tulajdonképpen hipergráfok helyett gráfok segítségével is számíthatók.

A 4. fejezetben  $m$ -multifakereső eljárásokat mutatunk be. Számos alkalmazásban – mint például a többdimenziós normális eloszlásfüggvény értékeire adott becslések – a metszetvalószínűségek kiértékelése költséges. Kritikus kérdés tehát, hogy vajon meg tudunk-e adni olyan algoritmust, mely csak „kevés” metszetvaló-

színűség kiértékelését teszi szükségessé. E szakaszban egyéb algoritmusok mellett látni fogunk olyan eljárást, mely mindössze a  $P(A_{k_1})$  ( $1 \leq k_1 \leq n$ ),  $P(A_{k_1} \cap A_{k_2})$  ( $1 \leq k_1 < k_2 \leq n$ ) valószínűségeket és  $\mathcal{O}(n)$   $P(A_{k_1} \cap \dots \cap A_{k_i})$  ( $1 \leq k_1 < \dots < k_i \leq n$ ,  $i = 3, \dots, m+1$ ) valószínűséget használ fel az összes  $\mathcal{O}(n^{m+1})$  közül. Az utolsó szakaszban az imént említett algoritmus tesztjeit mutatjuk be többdimenziós normális eloszlásfüggvények értékeire adott alsó becsléseken.

## 2. Multifák

Bevezetjük az  $m$ -multifa fogalmát. Rekurzív definíciót adunk a  $(V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$  hipergráfra, ahol  $V$  a csúcsok,  $\mathcal{E}_i$  pedig  $i$  csúcsból álló hiperélek halmaza.

1. *Definíció.* Legyen  $m$  egy pozitív egész szám. Az  $m$ -multicseresznye egy  $(V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$  alakú hipergráf, ahol  $V = \{v_1, \dots, v_{m+1}\}$  a csúcsok halmaza, és a hiperélek  $\mathcal{E}_i$  családja a  $V$  összes  $v_{m+1}$ -et tartalmazó  $i$ -csúcsú részhalmazaiából áll, azaz  $\mathcal{E}_i = \{H \mid v_{m+1} \in H \subset \{v_1, \dots, v_{m+1}\}, |H| = i\}$ . A  $v_{m+1}$  csúcsot az  $m$ -multicseresznye *domináns csúcsának* nevezzük. Az  $m$ -multicseresznyét, mely nem-domináns csúcsai a  $v_1, \dots, v_m$  és domináns csúcsa a  $v_{m+1}$ ,  $(\{v_1, \dots, v_m\}, v_{m+1})$ -gyel jelöljük.

Megjegyezzük, hogy egy 1-multicseresznye két csúcsból és az őket összekötő élből áll.

2. *Definíció.* Legyen  $m$  egy pozitív egész szám. Egy  $m$ -multifa egy  $(V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$  alakú hipergráf, ahol  $V$  a csúcsok halmaza,  $\mathcal{E}_i$  pedig  $i$ -csúcsú hiperélek családja. Egy  $m$ -multifát az alábbi szabályokkal leírt rekurzióval kaphatunk meg.

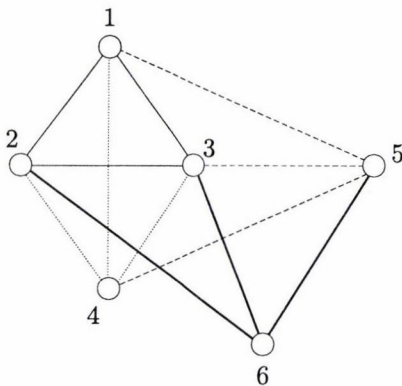
(i) A legkevesebb csúccsal rendelkező  $\Delta = (V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$   $m$ -multifának  $m$  csúcsa van,  $\mathcal{E}_i$  a  $V$  összes  $i$ -csúcsú részhalmazából áll (itt  $\mathcal{E}_{m+1} = \emptyset$ ).

(ii) Egy  $\Delta = (V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$   $m$ -multifából egy újabb  $\Delta' = (V', \mathcal{E}'_2, \dots, \mathcal{E}'_{m+1})$   $m$ -multifát nyerhetünk egy olyan  $(\{v_1, \dots, v_m\}, v_{m+1})$   $m$ -multicseresznye hozzávételével, melyre  $v_1, \dots, v_m \in V$  és  $v_{m+1}$  egy új csúcs (azaz  $v_{m+1} \notin V$ ). Tehát  $V' = V \cup \{v_{m+1}\}$ ,  $\mathcal{E}'_i = \mathcal{E}_i \cup \{H \mid v_{m+1} \in H \subset \{v_1, \dots, v_{m+1}\}, |H| = i\}$ .

Az 1. ábrán egy  $\Delta = (V, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4)$  3-multifát láthatunk, melyet az 1, 2, 3 csúcsokból és az  $\{1, 2\}$ ,  $\{1, 3\}$ ,  $\{2, 3\}$ ,  $\{1, 2, 3\}$  hiperélekből kiindulva rendre az  $(\{1, 2, 3\}, 4)$ ,  $(\{1, 3, 4\}, 5)$  és  $(\{2, 3, 5\}, 6)$  3-multicseresznyék hozzávételével kaphatunk meg. Az egyazon 3-multicseresznyékhez tartozó éleket ugyanazzal a stílusú vonallal rajzoltuk. A  $\Delta$  csúcsai és hiperélei tehát  $V = \{1, 2, 3, 4, 5, 6\}$ ,  $\mathcal{E}_2 = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}, \{2, 3\}, \{2, 4\}, \{2, 6\}, \{3, 4\}, \{3, 5\}, \{3, 6\}, \{4, 5\}, \{5, 6\}\}$ ,  $\mathcal{E}_3 = \{\{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\}, \{1, 3, 5\}, \{1, 4, 5\}, \{3, 4, 5\}, \{2, 3, 6\}, \{2, 5, 6\}, \{3, 5, 6\}\}$ ,  $\mathcal{E}_4 = \{\{1, 2, 3, 4\}, \{1, 3, 4, 5\}, \{2, 3, 5, 6\}\}$ .

Vegyük észre, hogy az 1-multifák a szokásos fák.

1. *Megjegyzés.* A  $\Delta = (V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$   $n$ -csúcsú  $m$ -multifára teljesül, hogy  $|\mathcal{E}_i| = \binom{m}{i} + \binom{m}{i-1}(n-m)$  bármely  $i = 2, \dots, m+1$  esetén.



1. ábra

3. *Definíció.* Legyenek  $A_1, \dots, A_n$  tetszőleges események. Ekkor a  $\Delta = (V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$   $m$ -multifa súlya:

$$w(\Delta) = \sum_{\{l_1, l_2\} \in \mathcal{E}_2} P(A_{l_1} \cap A_{l_2}) - \sum_{\{l_1, l_2, l_3\} \in \mathcal{E}_3} P(A_{l_1} \cap A_{l_2} \cap A_{l_3}) + \dots \\ \dots + (-1)^{m+1} \sum_{\{l_1, \dots, l_{m+1}\} \in \mathcal{E}_{m+1}} P(A_{l_1} \cap \dots \cap A_{l_{m+1}}).$$

1. *TÉTEL.* Legyenek  $A_1, \dots, A_n$  tetszőleges események és legyen  $\Delta = (V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$  egy tetszőleges  $m$ -multifa, melyre  $V = \{1, \dots, n\}$ . Ekkor

$$(1) \quad P(A_1 \cup \dots \cup A_n) \leq S_1 - w(\Delta),$$

ahol  $S_1 = \sum_{i=1}^n P(A_i)$ .

*Bizonyítás.* Az általánosság megszorítása nélkül feltehetjük, hogy  $\Delta$ -át a következő rekurzióval állítottuk elő. Kiindulunk a  $\Delta^{(m)} = (V^{(m)}, \mathcal{E}_2^{(m)}, \dots, \mathcal{E}_{m+1}^{(m)})$   $m$ -multifából, melyre  $V^{(m)} = \{1, \dots, m\}$ , valamint  $\mathcal{E}_i^{(m)}$  a  $V^{(m)}$  összes  $i$  elemű részhalmazából áll. Majd előállítjuk az  $m$ -multifáknak egy  $\Delta^{(m+1)}, \Delta^{(m+2)}, \dots, \Delta^{(n)}$  sorozatát a definíció értelmében oly módon, hogy  $j = m+1, \dots, n$  esetén  $\Delta^{(j-1)} = (V^{(j-1)}, \mathcal{E}_2^{(j-1)}, \dots, \mathcal{E}_{m+1}^{(j-1)})$ -ből a  $\Delta^{(j)} = (V^{(j)}, \mathcal{E}_2^{(j)}, \dots, \mathcal{E}_{m+1}^{(j)})$ -et az  $(\{i_1^j, \dots, i_m^j\}, j)$   $m$ -multicseresznye hozzávételével nyerjük, ahol  $1 \leq i_1^j < \dots < i_m^j \leq j-1$ . Végül  $\Delta = \Delta^{(n)}$ -nel kapjuk a tételben szereplő  $\Delta = (V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$   $m$ -multifát. A rekurzió alapján tehát

$$\mathcal{E}_{m+1} = (\{i_1^j, \dots, i_m^j, j\} \mid j = m+1, \dots, n),$$

valamint

$$\mathcal{E}_i = \bigcup_{j=m+1}^n \left\{ (K \mid j \in K \subset \{i_1^j, \dots, i_m^j, j\}, |K| = i) \right\} \cup \\ \bigcup \left\{ (K \mid K \subset \{1, \dots, m\}, |K| = i) \right\},$$

minden  $i = 2, \dots, m$  esetén.

$$\begin{aligned} P(A_1 \cup \dots \cup A_n) &= P(A_1 \cup \dots \cup A_{n-1}) + P(A_n) - P((A_1 \cup \dots \cup A_{n-1}) \cap A_n) = \\ &= P(A_1 \cup \dots \cup A_{n-2}) + P(A_{n-1}) - P((A_1 \cup \dots \cup A_{n-2}) \cap A_{n-1}) + P(A_n) - \\ &\quad - P((A_1 \cup \dots \cup A_{n-1}) \cap A_n) = \dots = S_1 - \sum_{j=2}^n P((A_1 \cup \dots \cup A_{j-1}) \cap A_j) = \\ &= S_1 - \left( \sum_{j=2}^m P((A_1 \cup \dots \cup A_{j-1}) \cap A_j) + \sum_{j=m+1}^n P((A_1 \cup \dots \cup A_{j-1}) \cap A_j) \right) \leq \\ &\leq S_1 - \left( \sum_{j=2}^m P((A_1 \cup \dots \cup A_{j-1}) \cap A_j) + \sum_{j=m+1}^n P((A_{i_1^j} \cup \dots \cup A_{i_m^j}) \cap A_j) \right) = \\ &= S_1 - \left( \sum_{j=2}^m P((A_1 \cap A_j) \cup \dots \cup (A_{j-1} \cap A_j)) + \right. \\ &\quad \left. + \sum_{j=m+1}^n P((A_{i_1^j} \cap A_j) \cup \dots \cup (A_{i_m^j} \cap A_j)) \right) \end{aligned}$$

Alkalmazva a szita formulát a  $P((A_1 \cap A_j) \cup \dots \cup (A_{j-1} \cap A_j))$  és  $P((A_{i_1^j} \cap A_j) \cup \dots \cup (A_{i_m^j} \cap A_j))$  kifejezésekre:

$$\begin{aligned} &S_1 - \left( \sum_{j=2}^m P((A_1 \cap A_j) \cup \dots \cup (A_{j-1} \cap A_j)) + \right. \\ &\quad \left. + \sum_{j=m+1}^n P((A_{i_1^j} \cap A_j) \cup \dots \cup (A_{i_m^j} \cap A_j)) \right) = \\ &= S_1 - \sum_{j=2}^m \left[ \sum_{1 \leq h_1 \leq j-1} P(A_{h_1} \cap A_j) - \sum_{1 \leq h_1 < h_2 \leq j-1} P(A_{h_1} \cap A_{h_2} \cap A_j) + \dots \right] \end{aligned}$$

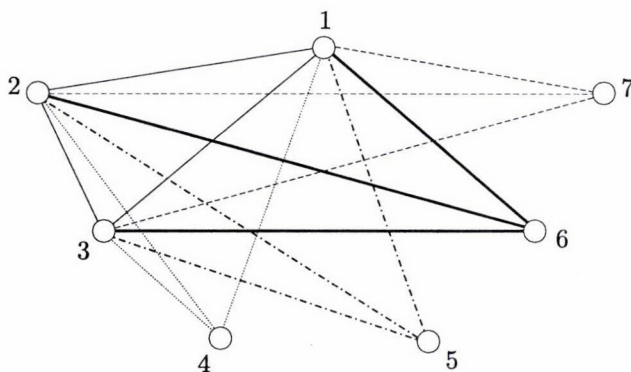
$$\begin{aligned}
& \dots + (-1)^{j-1} \sum_{1 \leq h_1 < \dots < h_{j-2} \leq j-1} P(A_{h_1} \cap \dots \cap A_{h_{j-2}} \cap A_j) + \\
& \quad + (-1)^j P(A_1 \cap \dots \cap A_{j-1} \cap A_j) \Big] - \\
& - \sum_{j=m+1}^n \left[ \sum_{\{k_1\} \subset \{i_1^j, \dots, i_m^j\}} P(A_{k_1} \cap A_j) - \sum_{\{k_1, k_2\} \subset \{i_1^j, \dots, i_m^j\}} P(A_{k_1} \cap A_{k_2} \cap A_j) + \dots \right. \\
& \quad \dots + (-1)^m \sum_{\{k_1, \dots, k_{m-1}\} \subset \{i_1^j, \dots, i_m^j\}} P(A_{k_1} \cap \dots \cap A_{k_{m-2}} \cap A_j) + \\
& \quad \left. + (-1)^{m+1} P(A_{i_1^j} \cap \dots \cap A_{i_m^j} \cap A_j) \right] = \\
& = S_1 - \sum_{\{l_1, l_2\} \in \mathcal{E}_2} P(A_{l_1} \cap A_{l_2}) + \sum_{\{l_1, l_2, l_3\} \in \mathcal{E}_3} P(A_{l_1} \cap A_{l_2} \cap A_{l_3}) - \dots \\
& \quad \dots + (-1)^m \sum_{\{l_1, \dots, l_{m+1}\} \in \mathcal{E}_{m+1}} P(A_{l_1} \cap \dots \cap A_{l_{m+1}}) = S_1 - w(\Delta). \quad \square
\end{aligned}$$

Az  $m = 1$  esetén a  $\Delta = (V, \mathcal{E}_2)$  maximális súlyú fával a Hunter–Worsley korlátot,  $m = 2$  esetén pedig a cseresznye fa korlátot kapjuk vissza.

**Numerikus példa:** Legyenek  $A_1, A_2, A_3, A_4, A_5, A_6, A_7$  a következő valószínűségekkel rendelkező események:

$$\begin{aligned}
p_1 &= p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = 0.35 \\
p_{1,2} &= 0.225, \quad p_{1,3} = 0.224, \quad p_{1,4} = 0.223, \quad p_{1,5} = 0.222, \quad p_{1,6} = 0.221, \\
p_{1,7} &= 0.220, \quad p_{2,3} = 0.219, \quad p_{2,4} = 0.218, \quad p_{2,5} = 0.217, \quad p_{2,6} = 0.216, \\
p_{2,7} &= 0.215, \quad p_{3,4} = 0.214, \quad p_{3,5} = 0.213, \quad p_{3,6} = 0.212, \quad p_{3,7} = 0.211, \\
p_{4,6} &= 0.210, \quad p_{4,6} = 0.209, \quad p_{4,7} = 0.208, \quad p_{5,6} = 0.207, \quad p_{5,7} = 0.206, \quad p_{6,7} = 0.205, \\
p_{1,2,3} &= p_{1,2,4} = \dots = p_{4,6,7} = p_{5,6,7} = 0.16 \\
p_{1,2,3,4} &= p_{1,2,3,5} = \dots = p_{3,5,6,7} = p_{4,5,6,7} = 0.13,
\end{aligned}$$

A 2. ábrán lévő 3-multifa alapján számított felső korlát 0.74. Ugyanahhoz a 3-multicseresznyéhez tartozó élek azonos vonalstílussal szerepelnek az ábrán.



2. ábra

Az alkalmazás során hasznunkra válhat az 1. tétel alábbi következménye.

1. KÖVETKEZMÉNY. Ha  $A_1, \dots, A_n$  tetszőleges események, akkor  $1 \leq m \leq n-1$  esetén  $P(A_1 \cup \dots \cup A_n) \leq \sum_{i=1}^{m+1} (-1)^{i+1} \left( \binom{m}{i} + \binom{m}{i-1}(n-m) \right) / \binom{n}{i} S_i$ , ahol  $S_i = \sum_{1 \leq k_1 < \dots < k_i \leq n} P(A_{k_1} \cap \dots \cap A_{k_i})$ .

*Bizonyítás.* Adjuk össze a tételben szereplő egyenlőtlenséget az összes olyan  $m$ -multifára, mely csúcsainak halmaza  $\{1, \dots, n\}$ . Ha ezen  $m$ -multifák száma  $X$ , akkor a következő egyenlőtlenséget kapjuk:

$$XP(A_1 \cup \dots \cup A_n) \leq \sum_{i=1}^{m+1} (-1)^{i+1} X \left( \left( \binom{m}{i} + \binom{m}{i-1}(n-m) \right) / \binom{n}{i} \right) S_i,$$

hiszen egy  $m$ -multifa  $i$  csúcsot tartalmazó éleinek száma  $\binom{m}{i} + \binom{m}{i-1}(n-m)$  bármely  $i = 2, \dots, m+1$  esetén; így  $X \left( \binom{m}{i} + \binom{m}{i-1}(n-m) \right) / \binom{n}{i}$ -szer szerepel egy adott  $i$ -csúcsú él az összegzésben, mivel mind az  $\binom{n}{i}$   $i$ -csúcsú élt ugyanannyiszor számoljuk. Az egyenlőtlenség mindkét oldalát  $X$ -szel elosztva a tételben szereplő egyenlőtlenséget kapjuk.  $\square$

Megjegyezzük, hogy  $m=1$  esetén az  $S_1, S_2$  alapján adható éles korlátot;  $m=2$  esetén a  $\frac{3S_3}{S_2} + 3 \geq n$  feltétel mellett az  $S_1, S_2, S_3$  alapján adható éles korlátot;  $m=3$  esetén a  $\frac{-12S_4 + 3(n-4)S_3 + (n-2)S_2}{(n-2)S_2 - 3S_3} + 3 \geq n$  feltétel mellett pedig az  $S_1, S_2, S_3, S_4$  alapján adható éles korlátot nyerjük a következményben szereplő egyenlőtlenségből. Az imént említett korlátok élességét Prékopa bizonyította (lásd [10]). Az  $m \geq 4$  esetén nem ismeretes zárt formula az  $S_1, \dots, S_{m+1}$  alapján adható éles felső korlátra.

2. TÉTEL. A  $\Delta = (V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$   $m$ -multifa a  $\Gamma = (V, \mathcal{E}_2)$  gráf által teljesen meghatározott.

*Bizonyítás.* Legyen  $\Delta = (V, \mathcal{E}_2, \mathcal{E}_3, \dots, \mathcal{E}_{m+1})$  egy  $m$ -multifa. Igazolnunk kell, hogy egy tetszőleges  $\Delta$ -áéval megegyező csúcs- és élhalmazú  $\bar{\Delta} = (V, \mathcal{E}_2, \bar{\mathcal{E}}_3, \dots,$



$\overline{\mathcal{E}_{m+1}}$ )  $m$ -multifára teljesül, hogy  $\overline{\mathcal{E}_i} = \mathcal{E}_i$  minden  $i = 3, \dots, m+1$ -re, azaz, hogy  $\overline{\Delta} = \Delta$ . Legyen  $1, \dots, n$  a csúcsok egy  $\Delta$ -t előállító rekurzió lépéseinek sorrendjében vett számozása, azaz a  $\Delta$ -t megkaphatjuk úgy, hogy kiindulunk az  $1, \dots, m$  csúcsokból valamint a hozzájuk tartozó hiperélekből, majd rendre hozzávesszük az  $m+1, \dots, n$  csúcsokat a megfelelő hiperélekkel. Hasonlóan legyen  $\overline{1}, \dots, \overline{n}$  a csúcsok egy  $\overline{\Delta}$ -t előállító rekurzió lépéseinek sorrendjében vett számozása. Feltehetjük, hogy  $|V| \geq m+2$ , hiszen  $|V| = m$  és  $|V| = m+1$  esetén  $\Delta$  és  $\overline{\Delta}$  hiperélei a  $V$  összes legalább kételemű részhalmazai, amikor viszont  $\overline{\Delta} = \Delta$ . A  $\Gamma = (V, \mathcal{E}_2)$  gráf minden csúcsának foka legalább  $m$ . Egy csúcs fokán a továbbiakban mindig a  $\Gamma$ -beli fokát értjük, és két csúcsot akkor mondunk szomszédosnak, ha  $\Gamma$ -beli éllel össze vannak kötve. A  $|V| \geq m+2$  feltétel miatt az  $1, \dots, m$  csúcsok közt legfeljebb egy  $m$ -edfokú van és ugyanez igaz az  $\overline{1}, \dots, \overline{m}$  csúcsokra is. Ha a  $\Gamma$  gráfnak van az  $1, \dots, m, \overline{1}, \dots, \overline{m}$  csúcsoktól különböző  $v$   $m$ -edfokú csúcsa, mely  $\Gamma$ -beli szomszédjainak halmaza  $U$ , akkor  $(U, v)$   $m$ -multicseresznye mind  $\Delta$ -ban mind pedig  $\overline{\Delta}$ -ban. Hagyjuk el  $\Delta$ -ból,  $\overline{\Delta}$ -ból és  $\Gamma$ -ból a  $v$  csúcsot a hozzátartozó hiperélekkel illetve élekkel. Folytassuk az eljárást addig, amíg  $\Gamma$ -ából származtatott gráfnak van az  $1, \dots, m, \overline{1}, \dots, \overline{m}$  csúcsoktól különböző  $m$ -edfokú csúcsa.

A fentiek alapján feltehető, hogy a  $\Gamma$  gráfnak nincs az  $1, \dots, m, \overline{1}, \dots, \overline{m}$  csúcsoktól különböző  $m$ -edfokú csúcsa. Az  $n$  és  $\overline{n}$  csúcsok foka azonban  $m$ , hiszen ezeket a csúcsokat a  $\Delta$ -t illetve a  $\overline{\Delta}$ -t előállító rekurzió utolsó lépésében vettük. A  $|V| \geq m+2$  feltétel miatt  $n \in \{\overline{1}, \dots, \overline{m}\}$  és  $\overline{n} \in \{1, \dots, m\}$ . Másrészt pedig az  $1, \dots, m$  és  $\overline{1}, \dots, \overline{m}$  csúcsok között legfeljebb egy-egy  $m$ -edfokú csúcs van, – hiszen ez már az  $m+1$  és  $m+2$  ( $\overline{m+1}$  és  $\overline{m+2}$ ) domináló csúcsú multicseresznyék hozzávétele után teljesül – azaz a  $\Gamma$  gráfnak pontosan két  $m$ -edfokú csúcsa van,  $n$  és  $\overline{n}$ .

(i) Azt állítjuk, hogy a  $\Delta$  bármely  $m$ -multicseresznyéjének nemdomináló csúcsai páronként szomszédosak.

Tegyük fel, hogy az előbbi állítás nem teljesül. Tetszőleges  $w \in \{m+1, \dots, n\}$  csúcsra jelöljük a  $\Delta$   $w$  domináló csúcsú  $m$ -multicseresznyéjét  $(X(w), w)$ -vel, ahol  $X(w)$  a nemdomináló csúcsok halmaza. Legyen  $(X(v), v)$  a  $\Delta$ -t előállító rekurzió szerint utoljára vett olyan  $m$ -multicseresznye, melyre teljesül, hogy  $X(v)$ -nek van két nem szomszédos csúcsa.

Megmutatjuk, hogy  $v < n$ . Tegyük fel ugyanis, hogy  $v = n$ . Tetszőleges  $w \in \{\overline{m+1}, \dots, \overline{n}\}$  csúcsra jelöljük a  $\overline{\Delta}$   $w$  domináló csúcsú  $m$ -multicseresznyéjét  $(\overline{X}(w), w)$ -vel. Az  $n \in \{\overline{1}, \dots, \overline{m}\} = \overline{X}(\overline{m+1})$  miatt az  $\{\overline{1}, \dots, \overline{m+1}\}$  elemei az  $n$  és szomszédjai, valamint az  $\{\overline{1}, \dots, \overline{m+1}\}$  halmazban lévő bármely két csúcs szomszédos. Mivel  $n$  egy  $m$ -edfokú csúcs, ezért szomszédos csúcsainak halmaza  $\{\overline{1}, \dots, \overline{m+1}\} \setminus \{n\}$ , ahonnan következik, hogy  $X(n) = \{\overline{1}, \dots, \overline{m+1}\} \setminus \{n\}$ . A  $v = n$  feltétel miatt azonban  $X(n)$ -ben nem lehet bármely két csúcs szomszédos. Az ellentmondásból adódik, hogy  $v < n$ .

Igazoljuk, hogy  $v \in X(v+1)$ . Tegyük fel ugyanis, hogy  $v \notin X(v+1)$ . Ekkor mind a  $v$  mind a  $v+1$  csúcs  $m$ -edfokú a  $\Delta$ -át előállító rekurzió azon lépésében, amikor az  $(X(v+1), v+1)$  multicseresznyét vesszük. Ekkor az  $\overline{n} \in \{1, \dots, m\}$  és

$m < v$  miatt  $\bar{n}$ ,  $v$  és  $v+1$  három különböző  $m$ -edfokú csúcs. Tehát legalább három  $m$ -edfokú csúcs van a rekurciónak abban a lépésében, amikor az  $(X(v+1), v+1)$  multicseresznyét vesszük. Ebből kifolyólag legalább három  $m$ -edfokú csúcs van a rekurzió további lépéseiben is, hiszen egy további lépésben hozzávett újabb csúcs már legfeljebb csak egy  $m$ -edfokú csúcsnak lehet a szomszédja, mert az  $m$ -edfokú csúcsok nem szomszédosak és feltevésünk szerint a továbbiakban vett minden multicseresznye bármely két nemdomináló csúcsa szomszédos. A  $\Gamma$  gráfnak azonban pontosan két  $m$ -edfokú csúcsa van. Ezen ellentmondásból kifolyólag  $v \in X(v+1)$ .

Igazoljuk, hogy  $v+1, \dots, n$  csúcsok minden szomszédja  $Z \cup Y$  halmazban van, ahol  $Z = X(v+1) \setminus \{v\}$  és  $Y = \{v, \dots, n\}$ . Az  $X(v+1) \subset X(v) \cup \{v\}$ , hiszen  $v \in X(v+1)$  miatt  $X(v+1)$   $v$  csúcstól különböző elemei  $v$  szomszédjai kell legyenek. A rekurzió ezen lépésében azonban csak az  $X(v)$ -beli csúcsok a  $v$  szomszédjai, a később vett csúcsok pedig már nem lehetnek  $X(v+1)$ -ben. Hasonlóan adódik, hogy  $v+k-1 \in X(v+k)$  minden  $k = 2, \dots, n-v$ -re, amiből induktíven kapjuk, hogy  $X(v+k) \subset X(v+k-1) \cup \{v+k-1\} \subset X(v+k-2) \cup \{v+k-2, v+k-1\} \subset \dots \subset X(v+1) \cup \{v+1, \dots, v+k-1\} = Z \cup \{v, v+1, \dots, v+k-1\}$ . Innen  $X(v+k) \subset Z \cup Y$  adódik minden  $k = 1, \dots, n-v$ -re. Következésképpen  $v+k$  ( $k = 1, \dots, n-v$ ) minden szomszédja valóban  $Z \cup Y$  halmazban van, hiszen a nem  $X(v+k)$ -beli szomszédok  $Y$ -ban vannak.

Megjegyezzük, hogy  $X(v+1) \subset X(v) \cup \{v\}$ -ből következik, hogy  $X(v) = Z \cup \{u\}$ , ahol  $u$  az  $X(v) \setminus Z$  egyetlen eleme. A  $v$  csúcs minden szomszédja az  $u$  kivételével szintén  $Z \cup Y$  halmazban van, hiszen a nem  $X(v)$ -beli szomszédok  $Y$ -ban vannak. A  $Z \subset X(v+1)$  miatt a  $Z$  bármely két csúcsa szomszédos. Az  $u$  tehát nem minden  $Z$ -beli csúccsal szomszédos, hiszen feltevésünk szerint  $X(v) = Z \cup \{u\}$ -nak van két nem szomszédos csúcsa.

Igazoljuk, hogy  $Z \cup Y = \{\bar{1}, \dots, \bar{n}\}$ . Tegyük fel, hogy az iménti állítás nem igaz. Mivel  $n \in \{\bar{1}, \dots, \bar{m}\} = \bar{X}(m+1)$ , ezért  $\{\bar{1}, \dots, \bar{m}+1\}$  elemei az  $n$  és szomszédjai. Tehát  $\{\bar{1}, \dots, \bar{m}+1\} \subset Z \cup Y$ , hiszen láttuk, hogy az  $n$  minden szomszédja  $Z \cup Y$ -beli. Legyen  $l$  a legkisebb olyan szám, melyre  $\bar{l} \notin Z \cup Y$ . Az  $\bar{X}(\bar{l}) \subset \{\bar{1}, \dots, \bar{l}-1\} \subset Z \cup Y$ . Azonban a  $v+1, \dots, n$  csúcsok minden szomszédja  $Z \cup Y$  halmazban van, így  $\bar{l} \notin Z \cup Y$  miatt  $\bar{X}(\bar{l}) \subset Z \cup \{v\}$ . Sőt  $|Z \cup \{v\}| = m$  miatt  $\bar{X}(\bar{l}) = Z \cup \{v\}$ . Tehát  $\bar{l}$  szomszédja  $v$ -nek és így  $\bar{l} \notin Z \cup Y$  miatt az  $\bar{l}$  csak az  $u$  lehet. Az  $u$  csúcsnak azonban nem minden  $Z$ -beli csúcsa szomszédja, ami  $\bar{X}(u) = Z \cup \{v\}$ -nak ellentmond. Tehát valóban  $Z \cup Y = \{\bar{1}, \dots, \bar{n}\} = V$ .

Azonban  $u \in V$ -re teljesül, hogy  $u \notin Z \cup Y$ , ami nyilvánvalóan ellentmond  $Z \cup Y = V$ -nek. Ez az ellentmondás igazolja (i)-beli állításunkat, azaz a  $\Delta$  bármely  $m$ -multicseresznyéjének nemdomináló csúcsai páronként szomszédosak. Természetesen ugyanez igaz  $\bar{\Delta}$ -ra is.

(ii) Azt állítjuk, hogy a  $H \subset V$  pontosan akkor hiperéle a  $\Delta$  ( $\bar{\Delta}$ )  $m$ -multifának, ha  $H$  bármely két csúcsa szomszédos.

Tegyük fel, hogy  $H \subset V$  hiperéle  $\Delta$ -ának. Ekkor vagy  $H \subset \{1, \dots, m\}$  vagy  $H \subset (X(v), v)$  valamely  $v \in V$ -re. Azonban az  $1, \dots, m$  csúcsok páronként szom-

szédosak és (i) miatt az  $X(v) \cup \{v\}$  halmaz bármely két csúcsa szintén szomszédos, tehát  $H$  bármely két csúcsa szomszédos.

Megfordítva, legyen  $H \subset V$  egy olyan halmaz, mely bármely két csúcsa szomszédos, és legyen  $h$  a  $H$  halmaznak a  $\Delta$ -t ( $\bar{\Delta}$ -t) előállító rekurzió szerint legutoljára vett csúcsa. A  $H \setminus h \subset X(h)$ , hiszen  $h$  a  $H \setminus h$  minden elemével össze van kötve már a rekurzió azon lépésében, amikor az  $(X(h), h)$  multicseresznyét vesszük. A  $H$  tehát valóban hiperéle  $\Delta$ -ának ( $\bar{\Delta}$ -nak).

A (ii)-ből adódik, hogy a  $\Delta$  és  $\bar{\Delta}$  hiperéleinek halmaza megegyezik, amivel a tétel állítását igazoltuk.  $\square$

4. *Definíció.* A  $(V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$   $m$ -multifa gráfján a  $(V, \mathcal{E}_2)$  gráfot értjük.

Tehát egy  $m$ -multifát megadhatjuk a gráfjával. Könnyen eldönthető, hogy egy gráf valamely  $m$ -multifa gráfja-e vagy sem. Hagyjuk el ugyanis a gráf egy  $m$ -edfokú csúcsát a hozzátartozó élekkel, majd az így nyert gráf egy  $m$ -edfokú csúcsát a hozzátartozó élekkel és így tovább míg egy  $m$ -csúcsú teljes gráfot kapunk. Ha nem akadunk el az algoritmussal, azaz mindig tudunk  $m$ -edfokú csúcsot választani míg egy  $m$ -csúcsú teljes gráfhoz jutunk, akkor az eredeti gráf egy  $m$ -multifa gráfja, ha elakadunk, akkor pedig nem.

### 3. Az $m$ -multifákkal adott korlátok javítása

Ebben a szakaszban megvizsgáljuk, hogyan lehet egy  $m$ -multifát  $(m+1)$ -multifává kiegészíteni, és ezáltal a valószínűségi korláton javítani.

3. TÉTEL. Tetszőleges legalább  $(m+2)$ -csúcsú  $m$ -multifa  $(m+1)$ -multifává egészíthető ki, azaz ha  $\Delta = (V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$  egy  $m$ -multifa, akkor létezik olyan  $\Delta' = (V', \mathcal{E}'_2, \dots, \mathcal{E}'_{m+2})$   $(m+1)$ -multifa, melyre  $\mathcal{E}_i \subset \mathcal{E}'_i$  minden  $i = 2, \dots, m+1$  esetén. A  $\Delta'$  alapján adott felső korlát nem nagyobb, mint a  $\Delta$  alapján adott.

*Bizonyítás.* Az előző tétel bizonyításának elején leírtakhoz hasonlóan előállítjuk az  $m$ -multifák egy  $\Delta^{(m+1)}, \Delta^{(m+2)}, \dots, \Delta^{(n)}$  sorozatát, ahol  $\Delta = \Delta^{(n)}$ -nel kapjuk a tételben szereplő  $\Delta = (V, \mathcal{E}_2, \dots, \mathcal{E}_{m+1})$   $m$ -multifát. A tételben szereplő  $\Delta'$ -öt a következő rekurzióval állíthatjuk elő. Kiindulunk a  $\Delta^{(m+1)'} = (V^{(m+1)'}, \mathcal{E}_2^{(m+1)'}, \dots, \mathcal{E}_{m+2}^{(m+1)'})$   $(m+1)$ -multifából, ahol  $V^{(m+1)'} = \{1, \dots, m+1\}$ , valamint  $\mathcal{E}_i^{(m+1)'}$  a  $V^{(m+1)'}$  összes  $i$  elemű részhalmazaiából áll. Nyilvánvaló, hogy  $\mathcal{E}_i^{(m+1)} \subset \mathcal{E}_i^{(m+1)'}$  minden  $i = 2, \dots, m+1$ -re. Előállítjuk az  $(m+1)$ -multifáknak egy  $\Delta^{(m+2)'}, \Delta^{(m+3)'}, \dots, \Delta^{(n)'}$  sorozatát oly módon, hogy  $j = m+2, \dots, n$  esetén  $\Delta^{(j-1)'} = (V^{(j-1)'}, \mathcal{E}_2^{(j-1)'}, \dots, \mathcal{E}_{m+2}^{(j-1)'})$ -ből a  $\Delta^{(j)'} = (V^{(j)'}, \mathcal{E}_2^{(j)'}, \dots, \mathcal{E}_{m+2}^{(j)'})$ -et a  $(\{i_1^j, \dots, i_m^j, i_{m+1}^j\}, j)$   $(m+1)$ -multicseresznye hozzávételével nyerjük, ahol  $i_{m+1}^j$  az  $\{1, \dots, j-1\} / \{i_1^j, \dots, i_m^j\}$  halmaz tetszőleges eleme. Látható, hogy azokat a hiperéleket is hozzáveszük  $\mathcal{E}_i^{(j-1)'}$ -hez, amelyeket  $\mathcal{E}_i^{(j-1)}$ -hez vettünk hozzá a  $\Delta$

rekurzív előállításánál, tehát  $\mathcal{E}_i^{(j)} \subset \mathcal{E}_i^{(j)'}$  minden  $i = 2, \dots, m+1$ -re és minden  $j = m+1, \dots, n$ -re. Végül  $\Delta' = \Delta^{(n)'}$ -vel kapjuk a tételben szereplő  $\Delta' = (V', \mathcal{E}'_2, \dots, \mathcal{E}'_{m+2})$   $(m+1)$ -multifát, melyre az előző mondat értelmében  $\mathcal{E}_i \subset \mathcal{E}'_i$  teljesül minden  $i = 2, \dots, m+1$  esetén.

Igazolnunk kell még, hogy a  $\Delta'$  alapján adott felső korlát nem nagyobb, mint a  $\Delta$  alapján adott, azaz  $w(\Delta) \leq w(\Delta')$ . Ehhez elég látni, hogy a fenti  $\Delta^{(j)}$  és  $\Delta^{(j)'}$  multifákra teljesül, hogy  $w(\Delta^{(j)}) \leq w(\Delta^{(j)'})$  minden  $j = m+1, \dots, n$  esetén. A  $j$ -re való indukcióval bizonyítunk. Ha  $j = m+1$ , akkor  $V^{(m+1)} = V^{(m+1)'}$ ,  $\mathcal{E}_i^{(m+1)} = \mathcal{E}_i^{(m+1)'}$  minden  $i = 2, \dots, m+1$  esetén és  $\mathcal{E}_{m+2}^{(m+1)} = \emptyset$ , következésképpen  $w(\Delta^{(m+1)}) = w(\Delta^{(m+1)'})$ . Rögzítsük  $j$ -t ( $m+2 < j < n$ ) és tegyük fel, hogy  $w(\Delta^{(j-1)}) \leq w(\Delta^{(j-1)'})$  teljesül. Be kell látnunk, hogy  $w(\Delta^{(j)}) \leq w(\Delta^{(j)'})$ . Ehhez elég igazolni, hogy

$$w(\Delta^{(j)}) - w(\Delta^{(j-1)}) \leq w(\Delta^{(j)'}) - w(\Delta^{(j-1)'}),$$

azaz

$$\begin{aligned} & \sum_{\substack{\{l_1, l_2\} \subset \\ \mathcal{E}_2^{(j)} \setminus \mathcal{E}_2^{(j-1)}}} P(A_{l_1} \cap A_{l_2}) - \dots + (-1)^{m+1} \sum_{\substack{\{l_1, \dots, l_{m+1}\} \subset \\ \mathcal{E}_{m+1}^{(j)} \setminus \mathcal{E}_{m+1}^{(j-1)}}} P(A_{l_1} \cap \dots \cap A_{l_{m+1}}) \leq \\ & \leq \sum_{\substack{\{l_1, l_2\} \subset \\ \mathcal{E}_2^{(j)'} \setminus \mathcal{E}_2^{(j-1)'}}} P(A_{l_1} \cap A_{l_2}) - \dots + (-1)^{m+2} \sum_{\substack{\{l_1, \dots, l_{m+2}\} \subset \\ \mathcal{E}_{m+2}^{(j)'} \setminus \mathcal{E}_{m+2}^{(j-1)'}}} P(A_{l_1} \cap \dots \cap A_{l_{m+2}}), \end{aligned}$$

mely a következő formába írható:

$$\begin{aligned} & \sum_{\substack{\{k_1\} \subset \\ \{i_1^j, \dots, i_m^j\}}} P(A_{k_1} \cap A_j) - \dots + (-1)^{m+1} \sum_{\substack{\{k_1, \dots, k_m\} \\ \subset \{i_1^j, \dots, i_m^j\}}} P(A_{k_1} \cap \dots \cap A_{k_m} \cap A_j) \leq \\ & \leq \sum_{\substack{\{k_1\} \subset \\ \{i_1^j, \dots, i_m^j, i_{m+1}^j\}}} P(A_{k_1} \cap A_j) - \dots + \\ & + (-1)^{m+2} \sum_{\substack{\{k_1, \dots, k_m\} \subset \\ \{i_1^j, \dots, i_m^j, i_{m+1}^j\}}} P(A_{k_1} \cap \dots \cap A_{k_{m+1}} \cap A_j) \end{aligned}$$

A bal oldalt mindkét oldalból kivonva a következő egyenlőtlenség kapjuk:

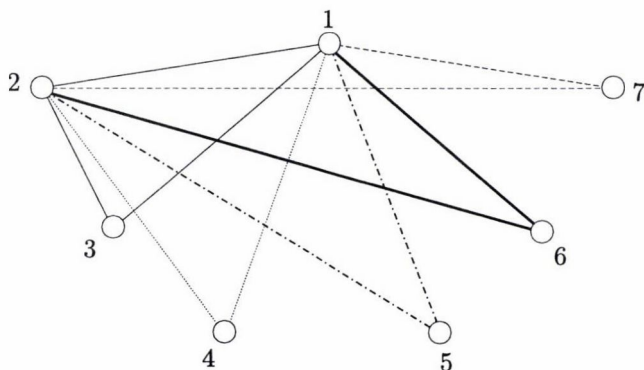
$$0 \leq P(A_{i_{m+1}^j} \cap A_j) - \sum_{\{h_1\} \subset \{i_1^j, \dots, i_m^j\}} P(A_{h_1} \cap A_{i_{m+1}^j} \cap A_j) + \dots$$

$$\begin{aligned}
& \dots + (-1)^{m+1} \sum_{\{h_1, \dots, h_{m-1}\} \subset \{i_1^j, \dots, i_m^j\}} P(A_{h_1} \cap \dots \cap A_{h_{m-1}} \cap A_{i_{m+1}^j} \cap A_j) + \\
& + (-1)^{m+2} P(A_{i_1^j} \cap \dots \cap A_{i_{m+1}^j} \cap A_j) = \\
& = P(A_{i_{m+1}^j} \cap A_j) - P((A_{i_{m+1}^j} \cap A_j) \cap (A_{i_1^j} \cup \dots \cup A_{i_m^j})),
\end{aligned}$$

a szita formula következtében. Ez utóbbi egyenlőtlenség pedig triviálisan teljesül.

□

A Hunter–Worsley korláton tehát javíthatunk, ha a maximális súlyú (1-multi)-fát  $m$ -multifává egészítjük ki a fenti tétel bizonyításában leírtak alapján. Az újabb korlát számítása persze több metszetvalószínűség ismeretét igényli. Tekintsük a fenti numerikus példát. Az 1 csúcsból kiinduló élek a maximális súlyú fa élei, mely alapján a Hunter–Worsley korlát 1.115-nek adódik. A maximális súlyú (1-multi)fa a 3. ábrán lévő 2-multifává (cseresznyefává) egészíthető ki, mely alapján számított korlát 0.83. A korlát további javítását nyerjük, ha a 2-multifát a 2. ábrán lévő 3-multifává egészítjük ki. Látható, hogy az (1-multi)fát 3-multifává, általában pedig egy  $m$ -multifát egy  $(m+r)$ -multifává közvetlenül is kiegészíthetünk a fenti tétel bizonyításában leírtakhoz hasonlóan. A  $\Delta^{(m+r)'} = (V^{(m+r)'}, \mathcal{E}_2^{(m+r)'}, \dots, \mathcal{E}_{m+2}^{(m+r)'})$   $(m+r)$ -multifából indulunk ki, ahol  $V^{(m+r)'} = \{1, \dots, m+r\}$ , így  $\Delta^{(m+r)'}$  egyértelmű. A  $\Delta^{(j-1)'}$ -ből a  $\Delta^{(j)'}-\text{öt}$  az  $(\{i_1^j, \dots, i_m^j, i_{m+1}^j, \dots, i_{m+r}^j\}, j)$   $(m+r)$ -multicseresznye hozzávételével nyerjük, ahol  $\{i_{m+1}^j, \dots, i_{m+r}^j\}$  az  $\{1, \dots, j-1\} / \{i_1^j, \dots, i_m^j\}$  halmaz tetszőleges részhalmaza és  $m+r < j \leq n$ .

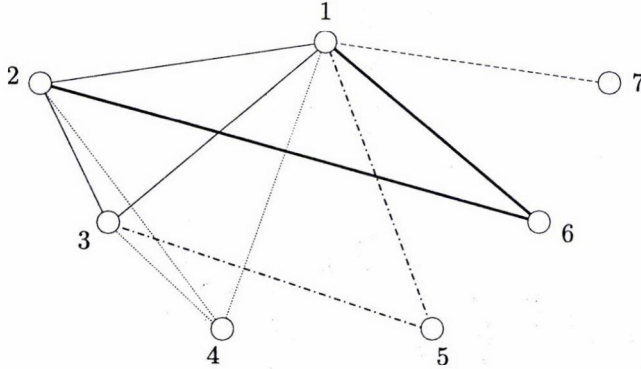


3. ábra

Hasonlóan látható, hogy egy  $m$ -multifa részlegesen is kiegészíthető. Példaként tekintsük a fenti numerikus példában szereplő maximális súlyú fa 4. ábrán lévő hipergráffá való kiegészítését. A fához rendre az  $(\{1, 2\}, 3)$ ,  $(\{1, 2, 3\}, 4)$ ,  $(\{1, 3\}, 5)$

és  $(\{1, 2\}, 6)$  multicseresznyéket vettük hozzá. A 4. ábrán lévő hipergráf a 2. ábrán lévő 3-multifa részhipergráfja, az általa kapott korlát

Hunter–Worsley korlát  $-(p_{2,3} + p_{2,4} + p_{2,6} + p_{3,4} + p_{3,5}) + (p_{1,2,3} + p_{1,2,4} + p_{1,2,6} + p_{1,3,4} + p_{1,3,5} + p_{2,3,4}) - p_{1,2,3,4} = 0.865$ .



4. ábra

#### 4. Multifakereső algoritmusok

Számos alkalmazás esetén, mint amilyen például a többdimenziós eloszlásfüggvények értékeire valamint hálózatok megbízhatóságára adott korlátok számítása, a felhasználandó  $P(A_{k_1} \cap \dots \cap A_{k_i})$  valószínűségeket ki kell értékelnünk. A kiértékelés rendszerint igen drága és a korlát számításához szükséges idő jelentős része fordítódik rá. Kritikus kérdés tehát, hogy nyerhetők-e, „jó” korlátok pusztán „néhány”  $P(A_{k_1} \cap \dots \cap A_{k_i})$  valószínűség kiértékelése árán.

A következő  $m$ -multifakereső algoritmus a  $P(A_{k_1})$  ( $1 \leq k_1 \leq n$ ),  $P(A_{k_1} \cap A_{k_2})$  ( $1 \leq k_1 < k_2 \leq n$ ) valószínűségeken és az összesen  $\mathcal{O}(n^{m+1})$   $P(A_{k_1} \cap \dots \cap A_{k_i})$  ( $1 \leq k_1 < \dots < k_i \leq n$ ,  $i = 3, \dots, m+1$ ) valószínűségekből mindössze  $\mathcal{O}(n)$ -en alapszik. Az 1. tételből kitűnik, hogy célunk nagy súlyú multifa keresése. Legyen  $\Gamma$  egy teljes gráf, mely csúcsainak halmaza  $\{1, \dots, n\}$ . A  $\Gamma$  élein bevezetünk egy súlyfüggvényt oly módon, hogy a  $j$  és  $k$  csúcsokat összekötő él súlya  $P(A_j \cap A_k)$  legyen. Egy mohó algoritmussal nyert maximális súlyú (1-multi)fát  $m$ -multifává egészítünk ki a következő rekurzióval. A rekurzió egy általános lépése, mely egy  $r$ -multifát  $(r+1)$ -multifává egészít ki ( $r = 1, \dots, m-1$ ), a következő allépésekből áll. Az első allépésben vesszük azt az  $r+1$  csúcsot, melyet az  $r$ -multifa előállításakor először (az első két allépésben) vettünk. Ezen  $r+1$  csúcs összes részhalmozásával együtt alkotja a kiinduló  $(r+1)$ -multifánkat. Minden további allépésben azt az  $(r+1)$ -multicseresznyét vesszük hozzá egy már meglévő  $(r+1)$ -multifához, mely az  $r$ -multifa  $r$ -multicseresznyéjének a lehető legsúlyosabb éllel való kiegészítése. A 3. tételből következik, hogy a rekurzió bármely lépésében konstruált



multifa alapján számított korlát legalább olyan jó, mint az előző lépések alapján konstruált multifa alapján számított. A rekurzióval nyert  $m$ -multifa alapján számított korlát kiszámításához annyi  $P(A_{k_1} \cap \dots \cap A_{k_i})$  ( $1 \leq k_1 < \dots < k_i \leq n$ ,  $i = 3, \dots, m+1$ ) valószínűséget kell kiértékelnünk, ahány  $i$ -csúcsú ( $i \geq 3$ ) hiperéle van egy  $m$ -multifának, azaz  $\binom{m}{i} + \binom{m}{i-1}(n-m)$ -et. Ezt az algoritmust használtuk a következő szakasz numerikus tesztjeinél valamint a 2. és 3. ábrán látható 2-multifa illetve 3-multifa konstruálásához.

Sajnos  $m > 1$  esetén mohó algoritmussal nem kapható meg a maximális súlyú  $m$ -multifa (lásd [3]). Ráadásul sok alkalmazásnál a fenti algoritmussal ha az  $m$  értékét eggyel növeljük jobb korlátot kapunk, mint amelyet a maximális súlyú  $m$ -multifa szolgáltat.

## 5. Többdimenziós normális eloszlásfüggvény értékeire adott alsó korlátok tesztjei

Legyen  $(\xi_1, \dots, \xi_n)$  egy standard normális eloszlású vektorváltozó, mely eloszlásfüggvényét jelöljük  $F(x_1, \dots, x_n)$ -nel. Bevezetve az  $A_i = P(\xi_i < x_i)$  ( $i = 1, \dots, n$ ) jelölést kapjuk, hogy

$$F(x_1, \dots, x_n) = P(\xi_1 < x_1, \dots, \xi_n < x_n) = P(A_1 \cap \dots \cap A_n) = 1 - P(\bar{A}_1 \cup \dots \cup \bar{A}_n).$$

A  $P(\bar{A}_1 \cup \dots \cup \bar{A}_n)$ -re adott felső korlátokkal tehát alsó korlátokat nyerünk az eloszlásfüggvény értékeire. A szükséges  $P(\bar{A}_k)$  ( $1 \leq k \leq n$ ),  $P(\bar{A}_{k_1} \cap \bar{A}_{k_2})$  ( $1 \leq k_1 < k_2 \leq n$ ) valószínűségeket az IMSL könyvtár mdnorr illetve mdnorr szubrutinjával, a  $P(\bar{A}_{k_1} \cap \bar{A}_{k_2} \cap \bar{A}_{k_3})$  ( $1 \leq k_1 < k_2 < k_3 \leq n$ ) értékeket Schervisch Fortran kódjával a NORINT7-tel [16]  $10^{-6}$  pontossággal, a  $P(\bar{A}_{k_1} \cap \dots \cap \bar{A}_{k_i})$  ( $1 \leq k_1 < \dots < k_i \leq n$ ,  $i \geq 4$ ) valószínűségeket pedig Genz Fortran kódjával a SADMVN-nel [4]  $i = 4, 5$  esetén  $10^{-6}$  illetve  $i = 6, 7, 8$  esetén  $10^{-5}$  pontossággal számítottuk.

A korlátokat az előző szakaszban leírt módon kaptuk. Minden esetben Szántai módszerével [17] kiszámítottuk az eloszlásfüggvény becslt értékét is. Szántai algoritmus először a 2-multifa alsó és a Tomescu felső [20] korlátokat számítja ki, majd Monte Carlo szimulációt használva becslt értéket ad a két érték közt lévő eloszlásfüggvényértékekre. A szimulációnál százezres mintát használtunk. Minden teszthez ötven feladatot generáltunk. A táblázatokban a korlátok és a becslt érték közti átlagos különbségeket valamint a számítási időket tüntettük fel. A számításokat Pentium 200MHz számítógéppel végeztük.

**1. Teszt.** Az ötven 40-dimenziós eloszlásfüggvényhez tartozó korreláció-mátrixokat Marsaglia és Olkin [9] módszerével állítottuk elő. A módszer szerint először egy  $\hat{C}$  alsó háromszög mátrixot készítünk el, mely elemeit a  $[-1, 1]$  intervallumon egyenletes eloszlással generáljuk. A  $C$  mátrixot a  $\hat{C}$  sorainak 2-norma szerinti lenormálásával nyerjük, melyből a  $\Sigma = CC^t$  korreláció-mátrixot kapjuk. Az

$x_i$  koordinátákat a  $[2, 3]$  intervallumon egyenletes eloszlással egymástól függetlenül generáltuk.

	Átlagos különbség a korlát és a becült érték között	Átlagos idő (másodpercekben)
Hunter-Worsley	0.037702	0.06
2-multifa	0.025482	0.09
3-multifa	0.017672	0.18
4-multifa	0.012275	0.47
5-multifa	0.008310	1.23
6-multifa	0.005429	3.27
7-multifa	0.003156	8.88
Szántai módszere	–	32.65

Szántai módszerének átlagos szórása 0.0002202 volt.

2. **Teszt.** Itt ugyanazt a 35-dimenziós eloszlásfüggvényt vettük ötven véletlenül választott  $x$  vektorral. A korrelációk  $\text{Corr}(\xi_i, \xi_j) = 0.8$  voltak minden  $i \neq j$  esetén. Az  $x_i$  koordinátákat a  $[0, 3]$  intervallumon egyenletes eloszlással egymástól függetlenül generáltuk.

	Átlagos különbség a korlát és a becült érték között	Átlagos idő (másodpercekben)
Hunter Worsley	0.278910	0.05
2-multifa	0.073927	0.17
3-multifa	0.025060	1.53
4-multifa	0.009387	4.54
5-multifa	0.003691	12.41
Szántai módszer	–	28.80

Szántai módszerének átlagos szórása 0.0008524 volt.

## Irodalom

- [1] G. Boole, *Laws of Thought*, American reprint of 1854 edition, Dover (New York, 1854).
- [2] C. E. Bonferroni, *Teoria Statistica Delle Classi e Calcolo Delle Probabilità*, Volume in onore di Riccardo Dalla Volta, Università di Firenze, 1-62 (1937).
- [3] J. Bukszár and A. Prékopa, *Probability Bounds with Cherry Trees*, RUTCOR Research Report (1999), 04-99.
- [4] A. Genz, Numerical Computation of the Multivariate Normal Probabilities, *J. Comput. Graph. Stat.*, 1 (1992), 141–150.
- [5] T. Hailperin, Best Possible Inequalities for the Probability of a Logical Function of Events, *The American Monthly*, 72 (1965), 343–359.



- [6] D. Hunter, An Upper Bound for the Probability of a Union, *J. Appl. Prob.*, **13** (1976), 597–603.
- [7] S. Kounias and J. Marin, Best Linear Bonferroni Bounds, *SIAM J. Appl. Math.*, **30**(2) (1976), 307–323.
- [8] S. M. Kwerel, Bounds on the Probability of a Union and Intersection of  $m$  Events, *Advances of Applied Probability*, **7** (1975), 431–448.
- [9] G. Marsaglia and I. Olkin, Generating Correlation Matrices, *SIAM Journal of Scientific and Statistical Computing*, **5** (1984), 470–475.
- [10] A. Prékopa, *Stochastic Programming*, Kluwer Academic Publishers (Dordrecht, 1995).
- [11] A. Prékopa, B. Vizvári and G. Regős, *Lower and Upper Bounds on Probabilities of Boolean Functions of Events*, RUTCOR Research Report (1995), 36–95.
- [12] A. Prékopa, Boole-Bonferroni Inequalities and Linear Programming, *Operations Research*, **36** (1988), 145–162.
- [13] A. Prékopa, Sharp Bounds on Probabilities Using Linear Programming, *Operations Research*, **38** (1990), 227–239.
- [14] A. Prékopa, The Discrete Moment Problem and Linear Programming, *Discrete Applied Mathematics*, **27** (1990), 235–254.
- [15] Y. S. Sathe, M. Pradhan and S. P. Shah, Inequalities for the Probability of the Occurrence of at least  $m$  out of  $n$  Events, *Journal of Applied Probability*, **17** (1980), 1127–1132.
- [16] M. Schervish, Multivariate Normal Probabilities with Error Bound, *Applied Statistics*, **33** (1984), 81–87.
- [17] T. Szántai, *Improved Bounds and Simulation Procedures on the Value of the Multivariate Normal Probability Distribution Function*, Proceedings of the VII. International Conference on Stochastic Programming, Vancouver, Canada Aug. 8–16 (1998) (submitted).
- [18] T. Szántai, Evaluation of a Special Multivariate Gamma Distribution, *Mathematical Programming Study*, **27** (1986), 1–16.
- [19] T. Szántai and J. Bukszár, Hiperceresznye-fákkal adott valószínűségi korlátok, *Alkalmazott Matematikai Lapok*, **19** (1999), 69–85.
- [20] I. Tomescu, Hypertrees and Bonferroni Inequalities, *Journal of Combinatorial Theory*, Series B **41** (1986), 209–217.
- [21] L. Takács, On the Method of Inclusion and Exclusion, *J. Am. Math. Assoc.*, **62** (1967), 102–113.
- [22] K. J. Worsley, An Improved Bonferroni Inequality and Applications, *Biometrika*, **69** (1982), 297–302.

(Beérkezett: 1999. június 7.)

BUKSZÁR JÓZSEF, MISKOLCI EGYETEM  
MATEMATIKA INTÉZET, ANALÍZIS TANSZÉK  
EGYETEMVÁROS, 3515 MISKOLC  
E-mail: MATBUK@GOLD.UNI-MISKOLC.HU

## PROBABILITY BOUNDS GIVEN BY MULTITREES

JÓZSEF BUKSZÁR

Upper bounds are presented for the probability of the union of finite number of events. The bounds are based on few terms of probabilities of intersections of at most  $m + 1$  events. The bounds are computed by means of special hypergraphs called  $m$ -multitrees. The main theorem is a generalization of the results of D. Hunter, K. J. Worsley and a former result of A. Prékopa and the author. The efficiency of the new bounds are shown by tests for estimation of the values of multivariate normal distribution functions.

## A TALAJVÍZ-SZENNYEZŐDÉS MODELLEZÉSE ÉS SZÁMÍTÁSA NAGYKÁTA TÉRSÉGÉBEN

KÉRI GERZSON ÉS RAPCSÁK TAMÁS

Budapest

Nagykátai városára (és a Tápió-régió egész területére) jellemző, hogy megoldatlan a szennyvízkezelés és szennyvízelvezetés, a közműháló nyitottsága igen nagy, mivel a vezetékes ivóvíz-hálózat ugyan 95%-ban kiépült, azonban a csatornázás és szennyvíz elvezetés mindmáig várat magára. A város 1998 óta rendelkezik csatorna hálózattal, de a közintézeteken kívül csak a lakások 7%-a van bekötve. Ez a kedvezőtlen helyzet — tartós fennállása esetén — káros környezeti terhelést okozhat. Szennyeződésterjedési modell felállításával és megoldásával számításokat végeztünk arra vonatkozóan, hogy a talajvízzel szállított szennyeződések mennyi idő alatt érhetik el a környező védett mocsarak területét. Az általunk korábban végzett hasonló számításoktól ([2], [3]) eltérően most nem pontszerű, hanem diffúz szennyezőforrással kellett dolgoznunk, mivel a szennyeződés nagy területen elszórtan keletkezik. Egy ennek a körülménynek megfelelően átalakított modell és megoldási algoritmus szerint végeztük el az ismertetésre kerülő szennyeződésterjedési számításokat. Az alkalmazott szennyeződésterjedési modell megoldása arra az eredményre vezetett, hogy a Nagykátától mintegy 1700 méterre levő mocsaras területen kb. 3500 év eltelte után jelenik meg a talajvíz útján szállított szennyeződés.

### 1. A város szennyvíz- és csatorna-helyzete

Nagykátai várost és az egész környező térséget az jellemzi, hogy szinte teljesen megoldatlan a szennyvízkezelés és szennyvízelvezetés. A térségben a közműháló nyitottsága az országos átlaghoz viszonyítva igen nagy. A vezetékes ivóvíz-hálózat 95%-ban kiépült, azonban a csatornázás és szennyvíz elvezetés mindmáig megoldatlan. A város 1998 óta rendelkezik ugyan csatorna hálózattal, de a közintézeteken kívül csak a lakások 7%-a van bekötve (forrás: [5], [6]). További gond, hogy a területet talaj- és rétegvíz süllyedés jellemzi; a víz nitráttartalma a sekélyebb kutakban a határérték környékén van és számítani kell annak mélyebbre történő fokozatos továbbterjedésére is (forrás: [4]).

A kedvezőtlen csatorna-helyzet miatt a keletkező szennyvíz nagy része szabálytalanul kialakított szikkasztókban kerül elhelyezésre. Ez nemcsak az emésztőgödörökben történik, hanem sok esetben a már nem használt kutakban is, melyekbe

átengedik az emésztőgyödrök tartalmát. A szippantott szennyvíz nagy részét illegális lerakóhelyeken tisztítják.

A város szennyvíz-helyzete számszerűen a következő adatokkal jellemezhető: Kiindulási adatként a napi vízfelhasználást vehetjük alapul, ami évszaktól függően 1800–3000 m<sup>3</sup>/nap. Ebből, a 7%-os mértékű csatornázottságot figyelembe véve, a csatornába kerülő szennyvíz 250–300 m<sup>3</sup>/nap. A többi 1550–2700 m<sup>3</sup>/nap szennyvízmennyiség a város területén vagy annak közvetlen környezetében elszikkad. A várost és annak környezetét terhelő szennyvízmennyiség éves összege mintegy 766500 m<sup>3</sup>/év. (A napi átlagnak vett 2100 m<sup>3</sup>/nap mennyiséget az év napjainak számával beszorozva adódik ez az összeg.)

Normális körülmények között a kommunális szennyvíz által okozott szennyezés nagy része a talajban és a vízben élő mikroorganizmusok tevékenysége folytán egy idő múlva lebomlik. Állandó és nagy terhelés esetén — ami Nagykáta területén jelenleg megvalósul — ez a folyamat igen lassan, vagy egyáltalán nem megy végbe, ezért az idők folyamán jelentős mértékű, a mélyebben fekvő rétegvizeket és a város szomszédságában elhelyezkedő védett területeket is veszélyeztető szennyezés következhet be (lásd [4]).

Szennyeződésterjedési modell felállításával és megoldásával számításokat végeztünk arra vonatkozóan, hogy a talajvízzel szállított szennyeződések mennyi idő alatt érik el a környező védett mocsarak területét. Az alkalmazott szennyeződésterjedési modell megoldása arra az eredményre vezetett, hogy amennyiben nem történik javulás a szennyvíz elvezetésében és tisztításában, a Nagykátától mintegy 1700 méterre levő mocsaras területen kb. 3500 év eltelte után jelenik meg a talajvíz útján szállított szennyeződés. (Megjegyezzük azonban, hogy ez a hosszú időtartam nem feltétlenül jelenti azt, hogy a térségben keletkező, és a talajvízbe kerülő szennyvíz teljesen veszélytelen lenne a környezetre. Jóval veszélyesebb ugyanis a szennyeződés függőleges irányú terjedése, amely esetleg már 50 éven belül elérheti a vízáadó rétegeket. A függőleges irányú szennyeződésterjedéssel azonban ebben a cikkben nem kívántunk foglalkozni.)

## 2. Az alkalmazott szennyeződésterjedési modell ismertetése

Egy szennyeződésterjedési modell elsődleges célja annak felmérése, hogy a különböző szennyezőanyagok terjedése milyen potenciális vagy tényleges veszélyhelyzetet idézhet elő, ha valamely szennyezőanyag túl nagy koncentrációt ér el a környezetben. A modellezés részét képező feladat annak eldöntése is, hogy a modellezés mely szennyezőkomponensekre terjedjen ki. Ebben a cikkben az egyik legártalmasabb komponens, a nitrát talajvízben történő terjedésére végzett számítások eredményét mutatjuk be.

A talajvíz szennyeződése a legnagyobb veszélyt akkor jelenti, ha a szennyezett talajvíz elér valamilyen ivóvízbázist. Fontos még a felszíni vizek védelme is a talajvíz útján történő szennyeződéstől. Ezért elsősorban azt kell mindig megnézni,

hogyan a talajvíz áramlásának az irányába eső területeken vannak-e és milyen távolságban vannak ivóvízbázisok, ill. folyók, tavak, mocsarak. Nagykátán és közvetlen környékén a mérési adatok szerint a talajvíz domináns iránya nyugatról keleti (pontosabban ettől az iránytól kb.  $10^\circ$ -kal eltérően északkeleti) irány. Amit az ebbe az irányba eső területen védeni kell, az a Hajta-patak ágai között és mellett levő kiterjedt mocsaras terület. Ennek Nagykátához legközelebb eső része a Hajta-patak Göböljárási-ágának nyugati oldalán, néhány kis tanya környékén helyezkedik el; s legközelebbi pontja Nagykáta lakott területének keleti csücskétől kb. 1700 méterre van (lásd 1. ábra). A még fokozottabb védelem alatt álló, Farmos községtől északra fekvő Nagy-nádas jóval távolabb, kb. 4350 méterre esik Nagykátától, és ennél is lényegesebb, hogy délkeleti irányba esik, vagyis olyan irányba, amerre Nagykáta felől gyakorlatilag nem jut el a talajvízben levő szennyeződés.

A feladat leírására a

$$(2.1) \quad \frac{\partial C}{\partial t} + \frac{v}{R} \cdot \frac{\partial C}{\partial x} = \frac{\alpha_L v}{R} \cdot \frac{\partial^2 C}{\partial x^2} + \frac{\alpha_T v}{R} \cdot \frac{\partial^2 C}{\partial y^2} - \lambda C$$

kétdimenziós transzportegyenletet alkalmazzuk, ahol  $v$  a szivárgás sebessége a pórusokban,  $\alpha_L$  a longitudinális diszperzivitás,  $\alpha_T$  a transzverzális diszperzivitás,  $R$  a késleltetési tényező,  $\lambda$  a bomlási együttható,  $t$  az idő,  $x$  a szivárgás irányában meghatározott térkoordináta,  $y$  pedig a szivárgás irányára merőleges térkoordináta. (A konkrét számítások során azonban el kell végeznünk a koordináta-rendszer eltolását és forgatását a szennyezőforrás helyének és a talajvíz áramlási irányának megfelelően.)

A transzportmodellek, transzportegyenletek részletes levezetése [1]-ben megtalálható. A szerzők az általános háromdimenziós transzportegyenlet felállítását követően megfogalmazzák ennek különböző egy-, ill. kétdimenziós speciális eseteit, és tárgyalják azok megoldási módszereit.

A legegyszerűbb, analitikus megoldási módszer gondolatmenetét követve, tekintünk először azt az esetet, amikor  $M$  tömegű pillanatnyi szennyezés jön létre a  $t = 0$  időpontban az  $x = 0$ ,  $y = 0$  helyen. Ebben az esetben a megoldás:

$$(2.2) \quad C(x, y, t) = \frac{M}{4\pi m n_0 v t \sqrt{\alpha_L \alpha_T}} \cdot \exp \left( -\frac{\left(x - \frac{vt}{R}\right)^2}{\frac{4\alpha_L vt}{R}} - \frac{y^2}{\frac{4\alpha_T vt}{R}} \right) \cdot \exp(-\lambda t),$$

ahol  $v$ ,  $\alpha_T$ ,  $R$  és  $\lambda$  ugyanaz, mint az előbb,  $m$  a víztartó vastagsága,  $n_0$  a szabad hézagterfogó (a talaj hézagainak, üregeinek térfogataránya, dimenzió nélküli faktor).  $\alpha_L$  a longitudinális diszperzivitás. A számított  $C(x, y, t)$  az adott tér- és időkoordinátákhoz tartozó szennyezőanyag-koncentráció ( $\text{mg}/\text{m}^3$ ).

Amennyiben egy egyenletes mértékben ható szennyezőforrás huzamos ideig szennyezi a talajvizet, abban az esetben a megoldásra (2.2)-ből az alábbi integrál-

formula vezethető le:

$$\begin{aligned}
 (2.3) \quad C(x, y, t) &= \frac{\mu}{4\pi mn_0 v \sqrt{\alpha_L \alpha_T}} \int_0^t \frac{1}{t - \tau} \\
 &\cdot \exp \left( - \frac{\left( x - \frac{v(t - \tau)}{R} \right)^2}{\frac{4\alpha_L v(t - \tau)}{R}} - \frac{y^2}{\frac{4\alpha_T v(t - \tau)}{R}} \right) \exp(-\lambda(t - \tau)) d\tau \\
 &= \frac{\mu}{4\pi mn_0 v \sqrt{\alpha_L \alpha_T}} \int_0^t \varphi(x, y, t; \tau) d\tau
 \end{aligned}$$

ahol  $\mu$  az időegység alatt kibocsátott szennyeződés és

$$\begin{aligned}
 (2.4) \quad &\varphi(x, y, t; \tau) \\
 &= \frac{1}{t - \tau} \cdot \exp \left( - \frac{\left( x - \frac{v(t - \tau)}{R} \right)^2}{\frac{4\alpha_L v(t - \tau)}{R}} - \frac{y^2}{\frac{4\alpha_T v(t - \tau)}{R}} \right) \exp(-\lambda(t - \tau)).
 \end{aligned}$$

A feladat közelítő megoldásának kiszámítását a

$$\begin{aligned}
 (2.5) \quad C(x, y, t) &\approx \frac{\mu}{4\pi mn_0 v \sqrt{\alpha_L \alpha_T}} \sum_{i=1}^t \varphi \left( x, y, t; i - \frac{1}{2} \right) = \frac{\mu}{4\pi mn_0 v \sqrt{\alpha_L \alpha_T}} \\
 &\cdot \sum_{i=1}^t \frac{\exp \left( - \lambda \left( t - i + \frac{1}{2} \right) \right)}{t - i + \frac{1}{2}} \exp \left( - \frac{\left( x - \frac{v \left( t - i + \frac{1}{2} \right)}{R} \right)^2}{\frac{4\alpha_L v \left( t - i + \frac{1}{2} \right)}{R}} - \frac{y^2}{\frac{4\alpha_T v \left( t - i + \frac{1}{2} \right)}{R}} \right)
 \end{aligned}$$

formulával végezzük (integrálközelítés érintő-formulával).

A [2] cikkben egy Gyál külterületén létesítendő (időközben már átadott) kommunális hulladéklerakó környezeti tanulmányához kapcsolódóan végeztünk szennyeződésterjedési számításokat. A modellezéshez és a számításhoz pontszerű szennyezőforrást feltételeztünk, ami akkor jogos volt, ugyanis egy feltételezett műszaki hiba, a szigetelés megsérülése folytán fellépő talajvízszennyeződésre alkalmaztuk a terjedési modellt. Hasonlóan, egy másik hulladéklerakó (Taksony), továbbá egy

aluminiumkohó mellett kialakított salak-depó (Apc) és egy szennyvíziszap komposztáló üzem (Dunaharaszti) esetében is pontszerű szennyezőforrást feltételezve végeztük a számításokat, melyekről a [3] cikkben számolunk be.

A transzportmodellek pontos megoldásának és közelítő megoldásának a kapcsolatára, a közelítés hibájára, valamint a megoldások paraméterérzékenységére vonatkozó részletek ugyancsak a [3] cikkben találhatók.

Az általunk korábban végzett hasonló, talajvízre vonatkozó szennyeződésmodellezési számításoktól eltérően (lásd [2], [3]) most nem feltételezhetünk pontszerű szennyezőforrást, mivel a szennyeződés nagy területen, gyakorlatilag Nagykáta város teljes területén keletkezik. (Olyan részletekkel a modellezés során nem tudtunk foglalkozni, hogy Nagykátán melyek a csatornázott és melyek a csatornázatlan területek, ezért a város teljes beépített területét tekintettük egyetlen diffúz szennyezőforrásnak). A számításokat azonban egy, a város területére képzelt raszteren végeztük, így végül mégis alkalmazni tudtuk a pontszerű szennyezőforrás esetére vonatkozó formulákat. Az ilymódon felépített modell szerint számított szennyezőanyag-koncentráció az egyes raszterpontokhoz külön-külön számított koncentráció-adatok eredője (összege). A diffúz szennyeződésterjedési modell ilyen közelítő megoldásával adódó eredményt a 2. ábra mutatja. Az ábra egy meglehetősen hosszú időtartam — 3500 év — után bekövetkező, a modellünk szerinti, állapotot mutatja, mivel azt akarjuk vele illusztrálni, hogy a Nagykátától keletre-délkeletre elhelyezkedő mocsaras területet kb. ennyi idő alatt éri el a talajvíz útján terjedő szennyeződés. (A számítást magunknak elvégeztük és a hasonló térképet elkészítettük több különböző időtartamra.)

Összehasonlításként lefuttattuk a számításokat a korábban, más feladatokra alkalmazott, pontszerű szennyezőforrásra vonatkozó megoldási algoritmussal is. Ehhez először olyan pontszerű szennyezőforrást feltételeztünk, mintha a város belterületének középpontjában, a térképen csak kis pontnak látszó helyen, koncentráltan képződne az évi  $766500 \text{ m}^3$ -re becsült szennyvízmennyiség. Annak feltételezése, hogy a város ilyen kis részén képződhetne egy év alatt ilyen tetemes szennyvízmennyiség, eléggé durva ötlet, s ennek megfelelően nem kell csodálkoznunk, hogy a kapott eredmények is irreálisak (3. ábra). A különböző módon készített térképes ábrák összehasonlítása azonban sok szempontból tanulságos. (Erre később majd még visszatérünk.) Az ábrákról most röviden csak annyit említünk, hogy árnyalás jelzi a különböző koncentrációjú szennyeződések; minél sötétebb egy hely az alaptérkép (1. ábra) azonos pontjához viszonyítva, annál nagyobb a szennyeződés koncentrációja.

Az utóbbi számítást ezután újra elvégeztük azzal a változtatással, hogy a pontforrást áthelyeztük Nagykáta lakóterületének a keleti csücskéhez, mivel a szennyvízzel fertőzött talajvíz által veszélyeztetett legközelebbi terület a várostól keleti irányban fekszik (4. ábra).

A számítások technikai részleteit és eredményeit a következő szakaszban tárgyaljuk részletesebben.

### 3. A szennyeződésterjedési modell megoldása

A számítások során egyrészt metrikus, másrészt pixel adatokkal számolunk. A kettő közötti átszámítás az eljárásnak paramétere, melynek konkrét értéke a jelenlegi esetben: 1 pixel = 12,5 m. A folytonos (diffúz) szennyeződésterjedési modell diszkrét, pontszerű szennyezőforrásokkal való közelítéséhez alkalmazott raszter függőleges beosztását állandónak vesszük: a számítások során a függőleges lépésköz 5 pixel, vagyis 62,5 m. A raszter vízszintes beosztását azonban rugalmasan kezeljük, mivel célszerűnek látszott soronként — vagyis egy-egy szélességi körön — azonos számú raszterpontot adni meg a város lakott területére helyezve. Ahol egy szélességi kör hosszabb területet metsz ki a lakott terület idomjából, ott ritkábban, ahol viszont rövidebb területet metsz ki, ott sűrűbben helyezkednek el a raszterpontok. (Így a szó szoros értelmében már nem is raszterről van szó, inkább csak kvázi-raszternek nevezhetnénk a modellben alkalmazott pontrácsot.) A rácspontok elhelyezésének elvét az 5. ábrával illusztráljuk. Paraméterként lehet megadni, hogy soronként — vagyis egy-egy szélességi körön — hány raszterpont legyen. A soronként felvett raszterpontok száma az eljárás paramétere, melynek értékét a végleges számítás során 6-nak adtuk meg. Konkáv alapterület esetén (Nagykátai területe is ilyen) a vízszintes osztóvonalak több különálló darabot is kimetszhetnek a szóban forgó területből; ilyenkor az egyes kimetszett szakaszokat külön felosztjuk a paraméter értékének megfelelő számú raszterpontra, amint az 5. ábrán is látható. A raszterpontokhoz súlyokat rendelünk az őket tartó szakaszok hossza arányában, s a teljes évente „termelő” szennyeződés mennyiségét ( $\mu$ ) a súlyoknak megfelelően szétosztjuk a kijelölt raszterpontok között. A konkrét számítás során 59 sorban elhelyezkedő  $6 \times 59 = 354$  raszterponttal futtattuk az eljárás programját.

További paraméterként a szennyeződés terjedésének fő irányát és opcionálisan egy nyílásszöget is meg lehet adni, amely azokat az irányokat foglalja magában, amilyen irányokban a terjedés lehetséges. A Nagykatára vonatkozó konkrét számításnál a fő irány ÉK-i  $80^\circ$ , a nyílásszög  $30^\circ$  (lásd 6. ábra).

Paraméterként lehet még megadni 4 különböző határértéket, melyek alapján az eljárás programja által elkészített térkép az egyes határértékek túllépésének megfelelően 5 különböző színre színezi (vagy 4 különböző árnyalatú szürke színre árnyalja) a térképet. (Az árnyalásos változatot a nyomdai előállítás kedvéért dolgoztuk ki. Az eredetileg 5 szín azért csökkent így 4 árnyalatra, mivel a koncentrációban a legkisebb határérték alatti, de a számítás során még pozitív értékűnek talált helyekhez a színezés során halvány zöld színt rendeltünk, az árnyalás során viszont a megfelelő helyeken változatlanul hagytuk a térképet.)

Mielőtt rátérnénk a konkrét számításokhoz használt további adatok meghatározására, előbb még szólnunk kell néhány szót a porózus anyagban szivárgó talajvíz szivárgási sebességének meghatározásáról. Darcy törvénye szerint egy porózus közegben szivárgó folyadék sebességét a  $v_{sz} = k \cdot \frac{\Delta h}{l}$  formula szerint lehet kiszámítani, ahol  $v_{sz}$  a szivárgás sebessége,  $l$  a két, az áramlás irányában elhelyezkedő, figyelőkút távolsága,  $\Delta h$  pedig a bennük észlelt vízoszlop különbség.



A sebesség dimenziójú  $k$  szivárgási tényező értéke (más kifejezéssel hidraulikus vezetőképességnek is szokták nevezni) függ a porózus anyagtól és az áramló folyadéktól, valamint attól, hogy milyen irányban (vízszintes vagy függőleges) tekintjük a vezetőképességet. A laza üledékes kőzetek vízáteresztő képessége vízszintes irányban lényegesen jobb, mint függőlegesen. A mérések szerint a Nagykáta környékére jellemző vízszintes irányú szivárgási tényező értéke  $k = 3 \cdot 10^{-5}$  m/sec-nak vehető, a hidraulikus gradiens ( $\frac{\Delta h}{l}$ ) értéke pedig  $10^{-4}$ -nek. (Eszerint a talajvíz áramlásának síkjában a lejtés szöge nagyon csekély, kilométerenként mindössze 10 cm.) Ennek megfelelően  $v_{sz}$  értéke  $v_{sz} = 3 \cdot 10^{-9}$  m/sec.

A szennyeződésterjedés effektív sebességét ebből  $v_{eff} = \frac{v_{sz}}{n_0}$  alapján számíthatjuk, hiszen csak a szabad térfogatban történhet az áramlás. Mivel esetünkben  $n_0 = 0,25$ , ennek alapján  $v = v_{eff} = 12 \cdot 10^{-9}$  m/sec adódik, ami közelítőleg 0,38 m/év.

Most folytatjuk a számításokhoz használt adatok felsorolását:

- A talajvíz effektív szivárgási sebessége (a 2. szakasz képleteiben szereplő  $v$  konkrét értéke) a fentiek szerint 0,38 m/év.
- Az  $\alpha_L$  longitudinális diszperzivitást  $\alpha_L = 25$ , az  $\alpha_T$  transzverzális diszperzivitást pedig  $\alpha_T = 1,5$  értéknek vettük a számítás során.
- Az  $R$  késleltetési tényezőre a leggyakoribb  $R = 1,2$  értéket használtuk.
- A  $\lambda$  bomlási együttható értékét 0-nak vettük, vagyis a szennyezőanyag esetleges lebomlását elhanyagoltuk a számítás során.
- Az időegység (jelen esetben 1 év) alatt kibocsátott  $\mu$  szennyeződés mennyisége az évente keletkező 766500 m<sup>3</sup>/év szennyvíz figyelembevételével, és a keletkező szennyvízben köbméterenként 100 g nitrát szennyeződést feltételezve,  $\mu = 7665 \cdot 10^4$  g/év.
- A víztartó átlagos vastagsága kb.  $m = 5$  m.
- A szabad hézagtérfogat a fentebb,  $v$  kiszámításánál már szerepelt  $n_0 = 0,25$  érték.
- A szennyezőanyagkoncentrációt (nitrát szennyezőanyagra) különböző  $t$  időadatokra és különböző helyekre számítottuk ki. A számított  $C(x, y, t)$  szennyezőanyag-koncentráció (mg/m<sup>3</sup>) értéktartományait térképeken mutatjuk be.

Az eljárás programjának futtatásai során először azonos nagyságrendű limiteket alkalmazva (pl. 10, 20, 40, 80 g/m<sup>3</sup>), ezen limitek értékétől és a modellezett időtartamtól függetlenül mindig azt tapasztaltuk, hogy az ilyen limiteknek megfelelő, szennyeződéskoncentráció értelemben vett szintvonalak nagyon közel voltak egymáshoz. Ez egy meglepő változás a korábbi talajvíz szennyeződésterjedési számítási tapasztalatokhoz ([2], [3]) képest, és valószínűleg a nagyon kicsi áramlási sebesség miatt adódott ilyen eredmény. Ezért áttértünk egy 10-es kitevőjű exponenciális skálára, amelynél már kicsit jobban, bár nem jelentősen, széthúzódtak a határok.

Az eredményeket a 2. ábra térképe mutatja. A programot a korábban felsorolt bemeneti adatokkal és a következő limitekkel futtattuk: 1, 10, 100, 1000 g/m<sup>3</sup>. A megadott limiteknek megfelelően, a térképen az alábbi 4 fokozatú árnyalás jelent meg: Az 1000 g/m<sup>3</sup> fölötti nitrát koncentrációhoz tartozó helyek sötét tónusú

árnyalást, a 100 és 1000 g/m<sup>3</sup> közötti koncentrációhoz tartozók ennél egy fokkal világosabb, 10 és 100 g/m<sup>3</sup> között még egy fokkal világosabb. 1 és 10 g/m<sup>3</sup> között a leghalványabb fokozatú árnyalást kapták, míg az 1 g/m<sup>3</sup> alatti nitrát koncentrációhoz tartozó helyeket nem árnyaltuk. A folyamatosnak feltételezett szennyeződés különböző időtartamaira elvégezve a számításokat, azt találtuk, hogy a Nagykáta keleti oldali szélső házaitól 1700 méterre található mocsaras terület szélén kb. 3500 év eltelte után jelenik meg a talajvíz útján szállított nitrát szennyeződés, ha ez alatt az idő alatt nem változnának az állapotok. Ehhez a 3,5 évezred nagyságú időadathoz tartozó számítások térképen ábrázolt eredményét mutatjuk be az ábrán, s ugyanehhez az időadathoz tartoznak a pontszerű szennyezőforrásra vonatkozó algoritmussal kísérletképpen elvégzett számítások eredményei (3. és 4. ábra).

Mivel a térkép árnyalása során csak 4 különböző árnyalatot használtunk, az árnyalt térképek nem mutatják a különbséget az erősen szennyezett (azaz a térképen legsötétbben árnyalt) területek szennyeződési koncentrációja között. Ezt a hiányt most azzal pótoljuk, hogy megadjuk a modell számításai szerint legnagyobb és átlagos szennyeződési koncentrációt az erősen szennyezett összterületre szorítkozva, mindkét modell (diffúznak, ill. pontszerűnek tekintett szennyezőforrás) esetére:

• Modell (szennyezőforrás) jellege:	diffúz	pontszerű
• Térkép:	2. ábra	3. és 4. ábra
• A modell szerint erősen szennyezett terület nagysága	10,667 km <sup>2</sup>	0,595 km <sup>2</sup>
• Maximális koncentráció az erősen szennyezett területen:	56,5 kg/m <sup>3</sup>	6113 kg/m <sup>3</sup>
• Átlagos koncentráció az erősen szennyezett területen:	16,7 kg/m <sup>3</sup>	300 kg/m <sup>3</sup>
• Koncentráció szórása az erősen szennyezett területen:	10,2 kg/m <sup>3</sup>	543 kg/m <sup>3</sup>

Az eredmények néhány tanulsága:

1) A talajvíz-áramlás lassúsága miatt nem kell meglepődnünk, hogy a kapott eredmény szerint évezredekben mérhető az az idő, amíg a szennyeződés 2 kilométernél is kisebb távolságra eljut, ha ugyanis elosztjuk az 1700 méter távolságot a 0,38 m/év sebességgel, azt kapjuk, hogy a pórusokban áramló talajvíz ezt a távolságot több mint 4400 év alatt teszi meg. (A terjedési modell a pozitív értékű longitudinális diszperzivitás miatt eredményez ennél kisebb időtartamot: 3500 évet. A fizikai törvények szerint a szennyeződés gyorsabban eljut egy adott helyre, mint a szennyeződést szállító talajvíz.)

2) A nagyon lassú talajvíz áramlás miatt a kevésbé szennyezett területek felől az erősebben szennyezett területek felé (azaz vissza Nagykáta felé, ill. a szennyezőforrás felé) haladva a szennyeződés koncentrációja nagyon gyorsan növekszik. Az egyes határértékekhez tartozó szintvonalak sokkal közelebb vannak egymáshoz,

mint az eddigi alkalmazásaink során megszokott áramlási sebességek esetén. (Gyál és Apc körzetében a 1,5 m/nap adatot, Taksonynál az 1 m/nap adatot, Dunaharasztnál pedig a 0,2 m/nap adatot használtuk.)

3) A pontszerű szennyezőforrásra vonatkozó számítások eredménye Nagykáta esetén (3. és 4. ábra) több szempontból is irreális: Elsősorban amiatt, hogy a városnak és a környéknek csak nagyon kis részét mutatja szennyezettnek. Ezek az ábrák is tanulságosak azonban, mivel mutatják, hogy egy pontszerű szennyezőforrásból milyen szennyeződési csóva indul ki.

4) A pontszerű szennyezőforrást feltételező modell alkalmazásának irreális voltát mutatja az a tény is, hogy ebben az esetben az erősen szennyezett területen a koncentráció maximális és átlagos értéke is nagyságrendekkel meghaladja a diffúz szennyezőforrás esetére kidolgozott modell számítása szerinti hasonló értékeket. (Az előbbi modell számítása szerint adódó, köbméterenként több száz vagy több ezer kg nitrát szennyeződés halmozódása kifejezetten abszurd eredmény.)

5) Az előbbieket következtében az is érthető, hogy a városnak a mocsaras terület felőli szélén felvett pontszerű szennyezőforrás esetén az e modellhez tartozó szennyezőcsóva jobban benyúlik a mocsaras területre, mint a diffúz szennyeződésterjedési modell esetén. Centrálisan elhelyezett szennyezőforrást modellezve viszont a csóva ugyanannyi idő alatt a mocsaras területnek még a közelébe se nagyon ér el.

Fontos végül megemlíteni, hogy a vízszintes irányú szennyeződésterjedési számítások során kapott, a fentiekben ismertetett eredmények (a terjedés időbeli elhúzódása) nem jelentik azt, hogy a Nagykáta területén nagy mennyiségekben keletkező, és a talajvízbe kerülő szennyvíz teljesen veszélytelen lenne a környezetre. Függőleges irányú szivárgás során ugyanis ez a szennyeződés a mélyebb fekvésű rétegvizeket, és így az ivóvízbázist is veszélyezteti a számítások szerint már 50, 100 vagy 200 éven belül, a réteg mélységétől és a rétegsor összetételétől függően. A függőleges irányú szennyeződésterjedéssel azonban ebben a cikkben nem kívántunk foglalkozni.

### Köszönetnyilvánítás:

A szerzők köszönetüket fejezik ki Balla Katalinnak a kézirat átnézéséért, Bartha Béla Kolosnak a kézirat átnézéséért és a modell adatainak értelmezésében nyújtott segítségével. Orsovai Imrének a modell adatainak, paramétereinek a meghatározásában nyújtott segítségével és mindhármutknak hasznos tanácsaikért.

### Irodalom

- [1] Kovács B. és Szabó I., Hulladékellhelyezés IV. *A szennyezőanyagok terjedése. A modellezés elmélete és gyakorlata*, „Ipar a környezetért” alapítvány (1995).
- [2] Kéri G., Orsovai I. és Rapcsák T., Egy transzportmodell alkalmazása a Gyál térségében léte-sítendő hulladéklerakó esetleges talajszennyező hatásának vizsgálatára (Esettanulmány), *Alkalmazott Matematikai Lapok*, 20 (1999) 169-182.

- [3] Balla, K., Kéri, G. and Rapcsák, T., *Pollution of the underground water — a computational case study using a transport model*, to be published.
- [4] *Tápiómenti Kistérség Környezetvédelmi Programja*, PESTTERV (Budapest, 1999).
- [5] *Tápiómenti Kistérség Területfejlesztési Konceptiója*, PESTTERV (Budapest, 1999).
- [6] *Magyarország Megyei Kézikönyvei 13. (Pest Megye)* (1995), 613–623.

(Beérkezett: 1999. december 16.)

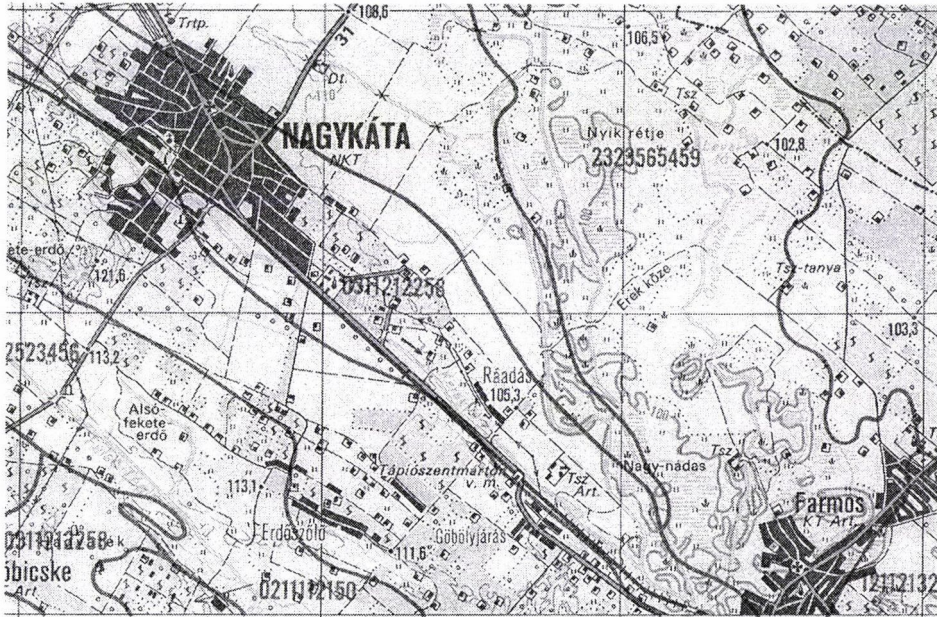
KÉRI GERZSON  
MTA SZTAKI  
BUDAPEST

RAPCSÁK TAMÁS  
MTA SZTAKI  
BUDAPEST

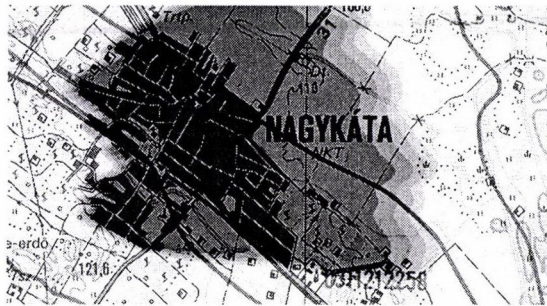
## MODELLING OF GROUND-WATER POLLUTION IN THE AREA OF NAGYKÁTA

GERZSON KÉRI AND TAMÁS RAPCSÁK

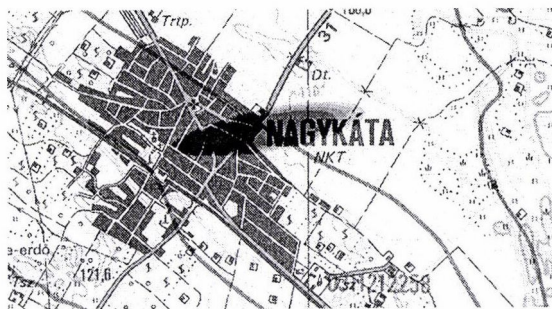
Ground water pollution is a real danger in the area of Nagykáta (i.e., the region of the river Tápió), because sewage handling and disposal is far from being solved in that area. Drainage system is built out only at a few settlement in the area, one of them is the town of Nagykáta (which is regarded the centre of the region), but even in Nagykáta only 7% of the houses are connected to the drainage system. For calculating the concentration of pollution in the ground water near the town, a transport model with a diffuse source of pollution has been applied.



1. ábra. Nagykáta és a tőle keletre és délre fekvő terület térképe



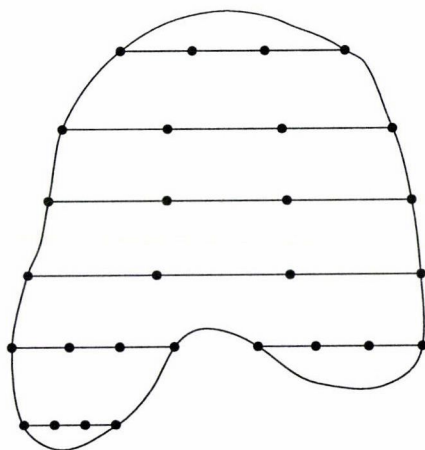
2. ábra. A talajvíz nitrát-szennyeződése Nagykáta környékén



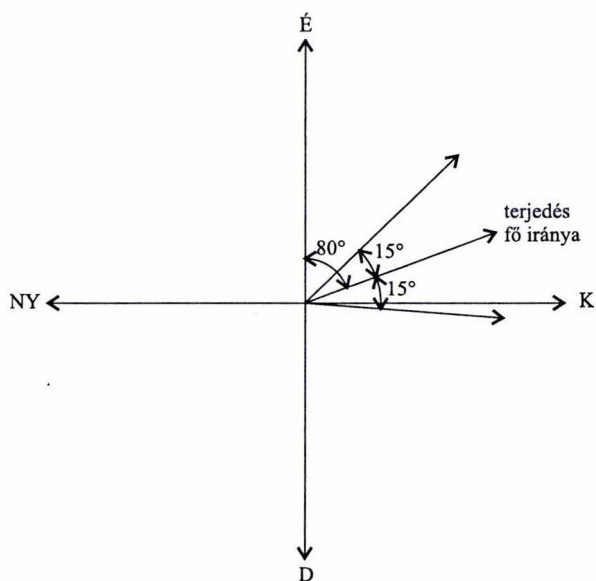
3. ábra. Nitrát szennyeződés pontszerű (centrális) szennyeződés feltételezésével



4. ábra. Nitrát szennyeződés pontszerű (nem centrális) szennyeződés feltételezésével



5. ábra. Raszterpontok kijelölésének módja diffúz szennyeződésterjedési modell alkalmazásakor



6. ábra. A szennyeződésterjedés iránya és nyílásszöge





## FELÜLETEK ELŐÁLLÍTÁSA TRIGONOMETRIKUSAN SÚLYOZOTT SPLINE-FÜGGVÉNYEK FELHASZNÁLÁSÁVAL

PETHŐNÉ VENDEL TERÉZ

Pécs

Térbeli pontokra illesztett trigonometrikusan súlyozott spline-görbét definiálunk, amelynek görbeívei görbület-folytonosan csatlakoznak egymáshoz, sőt a csatlakozópontoktól eltekintve végtelen sokszor differenciálhatók. Az ilyen görbével általános eltolási felületet hozunk létre. A felületeket egy vagy két trigonometrikusan súlyozott spline vezérgörbével definiáljuk, a generáló görbe pedig harmadfokú Bézier-görbe lesz. Így vegyes típusú felületfoltokat kapunk, amelyek normális-folytonosan csatlakoznak egymáshoz.

### 1. Bevezetés

Mivel háromdimenziós világunkban sok bonyolult alakú tárgy fordul elő, ezek megfelelő modellezésére és valósághű képi megjelenítésére a CAD-rendszerek spline-technikán alapuló matematikai módszereket használnak. Ebben a cikkben trigonometrikusan súlyozott spline-függvényekkel mutatunk be néhány megoldást görbe- és felületmodellezési problémákra.

A második fejezetben interpolációs pontokra illesztett görbét definiálunk trigonometrikusan súlyozott spline-ként, amelynek görbeíveit köríveknek és egyenes szakaszoknak konvex kombinációja állítja elő aszerint, hogy három egymás után következő pont egy körön vagy egy egyenesen van. Az így definiált spline-görbe görbület-folytonos lesz [3]. Alkalmazásként poligonok csúcsainak lekerekítésére mutatunk példát. A harmadik fejezetben olyan általános eltolási felületet definiálunk, amelynek vezérgörbéje trigonometrikusan súlyozott spline, generáló görbéje pedig egy harmadfokú Bézier-görbe. A felületfoltokat tehát vegyes típusú, egyik irányban trigonometrikus, a másik irányban pedig polinomiális kétparaméteres vektorfüggvény írja le. A szomszédos foltok  $GC^1$  rendben csatlakoznak egymáshoz. További alkalmazást mutatunk a negyedik fejezetben két vezérgörbével definiált szabadon formált eltolási felület származtatására és az alaki paraméterek megválasztására.

## 2. Interpolációs, trigonometrikusan súlyozott spline-görbe

A háromdimenziós  $\mathbf{E}^3$  euklideszi térben adott pontsorozaton áthaladó, azonos módon definiált ívekből összetett görbét készítünk úgy, hogy a görbe az ívek csatlakozási pontjaiban görbület-folytonos legyen. A görbéveket trigonometrikus függvényekkel fogjuk leírni, ezért azt mondjuk, hogy az adott pontsorozatra trigonometrikusan súlyozott interpolációs spline-görbét illesztünk. A két szomszédos pont közötti görbévre előírjuk, hogy egyenes szakaszt ill. körívet állítson elő abban az esetben, ha négy egymás utáni pont egy egyenesen ill. egy körön van.

### 2.1.

Legyen  $\{P_1, P_2, P_3, \dots, P_n\} \subset \mathbf{E}^3$  az adott interpolációs pontok halmaza. A görbét a következő algoritmus szerint definiáljuk [6]:

1. lépés: Az első kettő és az utolsó két pont közötti görbévek definiálása végett kiegészítjük a pontsorozatot a  $P_0$  és  $P_{n+1}$  pontokkal (lásd 2.2.).

2. lépés: Elkészítjük az  $i$ -edik görbévet ( $i = 1, \dots, n-1$ ):

– határozzunk meg egy kört vagy egyenest a  $P_{i-1}$ ,  $P_i$  és  $P_{i+1}$  pontokon keresztül, és ennek a  $P_i$  és  $P_{i+1}$  közötti darabja legyen ív  $[i]^{\text{bal}}$ .

– határozzunk meg egy kört vagy egyenest a  $P_i$ ,  $P_{i+1}$  és  $P_{i+2}$  pontokon keresztül, és ennek a  $P_i$  és  $P_{i+1}$  közötti darabja legyen ív  $[i]^{\text{jobb}}$ .

Az ív  $[i]^{\text{bal}}$  és ív  $[i]^{\text{jobb}}$  görbéveket paraméteres vektoregyenlettel adjuk meg a  $[0,1]$  intervallum felett. Az egyenes szakaszt az

$$(1-t) \cdot \mathbf{p}_i + t \cdot \mathbf{p}_{i+1}, \quad t \in [0,1]$$

vektorfüggvénnyel írjuk le, ahol  $\mathbf{p}_i$  a  $P_i$  pont,  $\mathbf{p}_{i+1}$  a  $P_{i+1}$  pont helyvektorát jelöli. A körívet, amelynek középpontja  $C$  és középponti szöge  $\alpha_i := \angle(P_i, C, P_{i+1})$ , a

$$\mathbf{c} + \mathbf{R}_i(t) \cdot (\mathbf{p}_i - \mathbf{c}), \quad t \in [0,1]$$

függvénnyel írjuk le, ahol  $\mathbf{c}$  a középpont helyvektora,

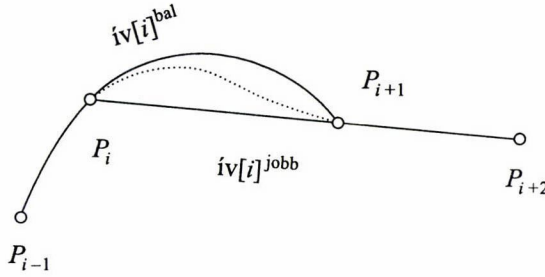
$$\mathbf{R}_i(t) = \begin{pmatrix} \cos(t\alpha_i) & -\sin(t\alpha_i) & 0 \\ \sin(t\alpha_i) & \cos(t\alpha_i) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

pedig a térbeli forgatás mátrixa a kör síkjában  $\alpha_i$  szögsebességgel egy megfelelő koordináta-rendszerben.

– képezzük az ív  $[i]^{\text{bal}}(t)$  és ív  $[i]^{\text{jobb}}(t)$  görbék konvex kombinációját úgy, hogy az eredményül kapott görbe olyan legyen  $P_i$  és  $P_{i+1}$  között, hogy a  $P_i$  ponthoz közel az ív  $[i]^{\text{bal}}(t)$ -re, a  $P_{i+1}$  környezetében pedig ív  $[i]^{\text{jobb}}(t)$ -re hasonlítson (1. ábra) [4]:

$$(1) \quad \text{gív}[i](t) := \cos^2\left(t \cdot \frac{\pi}{2}\right) \cdot \text{ív}[i]^{\text{bal}}(t) + \sin^2\left(t \cdot \frac{\pi}{2}\right) \cdot \text{ív}[i]^{\text{jobb}}(t),$$

$$t \in [0,1], \quad i = 1, \dots, n-1.$$

1. ábra. Az  $i$ -edik görbeív

3. lépés: Az ilyen módon meghatározott görbeívek uniója a kívánt interpolációs görbe.

Ez a görbe lokálisan módosítható, azaz alakja csak az elmozdított pont valamely jól meghatározható környezetében változik.

2.1. TÉTEL. A fenti (1) képlettel készített trigonometrikusan súlyozott spline-görbe ívei görbület-folytonosan csatlakoznak egymáshoz.

*Bizonyítás.* A görbületet az  $(i-1)$ -edik görbeív  $gív[i-1](t)$  és az  $i$ -edik görbedarab  $gív[i](\tau)$  ( $i = 2, \dots, n-1$ )  $t = 1$ , ill.  $\tau = 0$  csatlakozó pontjában vizsgáljuk. Tekintsük (1) deriváltjait:

$$\begin{aligned} \frac{d}{dt} gív[i-1](t) = & -\frac{\pi}{2} \cdot \sin(t \cdot \pi) \cdot ív[i-1]^{bal}(t) + \cos^2\left(t \cdot \frac{\pi}{2}\right) \cdot \frac{d}{dt} ív[i-1]^{bal}(t) + \\ & + \frac{\pi}{2} \cdot \sin(t \cdot \pi) \cdot ív[i-1]^{jobb}(t) + \sin^2\left(t \cdot \frac{\pi}{2}\right) \cdot \frac{d}{dt} ív[i-1]^{jobb}(t) \end{aligned}$$

és

$$\begin{aligned} \frac{d}{d\tau} gív[i](\tau) = & -\frac{\pi}{2} \cdot \sin(\tau \cdot \pi) \cdot ív[i]^{bal}(\tau) + \cos^2\left(\tau \cdot \frac{\pi}{2}\right) \cdot \frac{d}{d\tau} ív[i]^{bal}(\tau) + \\ & + \frac{\pi}{2} \cdot \sin(\tau \cdot \pi) \cdot ív[i]^{jobb}(\tau) + \sin^2\left(\tau \cdot \frac{\pi}{2}\right) \cdot \frac{d}{d\tau} ív[i]^{jobb}(\tau) \end{aligned}$$

$$t, \tau \in [0, 1], \quad i = 2, \dots, n-1.$$

A csatlakozási pontban

$$\begin{aligned} \frac{d}{dt} gív[i-1](1) &= \frac{d}{dt} ív[i-1]^{jobb}(1), \\ \frac{d}{d\tau} gív[i](0) &= \frac{d}{d\tau} ív[i]^{bal}(0) \end{aligned}$$

adódik. Hasonlóan a második deriváltakra

$$\frac{d^2}{dt^2} giv[i-1](1) = \frac{d^2}{dt^2} iv[i-1]^{jobb}(1)$$

$$\frac{d^2}{d\tau^2} giv[i](0) = \frac{d^2}{d\tau^2} iv[i]^{bal}(0)$$

teljesül.

Mivel az  $iv[i-1]^{jobb}(t)$  és az  $iv[i]^{bal}(\tau)$  görbéket ( $i = 2, \dots, n-1$ ) a  $P_{i-1}, P_i$  és  $P_{i+1}$  pontokon áthaladó csatlakozó  $\widehat{P_{i-1}P_i}$  és  $\widehat{P_iP_{i+1}}$  azonos sugarú körívek illetve egyenes szakaszok definiálják, ezért a másodrendű geometriai folytonosság teljesül, vagyis a görbéiek görbület-folytonosan csatlakoznak.

A  $P_{i-1}, P_i, P_{i+1}$  pontokon áthaladó ívnek a  $P_{i-1}$  és  $P_i$  közé eső darabja  $iv[i-1]^{jobb}(t)$ , a  $P_i$  és  $P_{i+1}$  pontok közé eső darabja pedig  $iv[i]^{bal}(\tau)$ . Bár mindkettő paramétertartománya a  $[0, 1]$  intervallum, a két ívhossz általában különböző. Ennek megfelelően  $\tau = \lambda \cdot t$ , ( $\lambda > 0$ ). Vegyük észre, hogy a fenti első deriváltak az ívhosszakkal fordítottan arányosak:

$$\frac{d}{dt} iv[i-1]^{jobb}(1) : \frac{d}{d\tau} iv[i]^{bal}(0) = \widehat{P_iP_{i+1}} : \widehat{P_{i-1}P_i}.$$

A fenti második deriváltak pedig az ívhosszak négyzetével fordítottan arányosak összhangban az ívhossz és a görbület szokásos képletével. Tehát a két görbe  $P_i$  pontbeli érintője párhuzamos, a görbületek pedig megegyeznek.  $\square$

Megjegyezzük, hogy körívek és egyenes szakaszok lineáris kombinációjával készített görbéiek csak érintő-folytonosan csatlakoznak [5, 6, 8].

## 2.2.

Az interpolációs görbe  $n-1$  görbeívből áll, és minden ív a két végpontjával és még két szomszédos ponttal meghatározott, átfedő egyenes ill. kör konvex kombinációja. Emiatt az első és utolsó görbeív definíciója megkívánja a pontsor kibővítését. A  $P_0$  és  $P_{n+1}$  pontok megválasztása különböző peremfeltételeknek megfelelően történhet. [8]

a) *Természetes peremfeltétel*

A kezdő- illetve a végpontban a görbe görbülete nulla. Ekkor

$$p_0 := p_1 + (p_1 - p_2),$$

$$p_{n+1} := p_n + (p_n - p_{n-1}),$$

ahol  $p_i$  a  $P_i$  pont helyvektorát jelöli.

b) *Periodikus peremfeltétel*

Ilyen esetben  $P_1 = P_n$ . Legyenek

$$P_0 := P_{n-1},$$

$$P_{n+1} := P_2.$$

c) *Körív szerinti peremfeltétel*

Mivel a 2.1. pontban definiált görbe egy körön elhelyezkedő négy pont esetén körívet ad vissza, a  $P_0$  ill.  $P_{n+1}$  pontokat az első három ill. az utolsó három pont által meghatározott körön vesszük fel. Ha a  $P_1, P_2, P_3$  pontokon átmenő kör középpontját  $C$ -vel, a  $\overrightarrow{CP_1}$  és  $\overrightarrow{CP_2}$  vektorok által bezárt szöget pedig  $\varphi$ -vel jelöljük, akkor a  $P_0$  pontot

$$p_0 := c + 2 \cdot \cos \varphi \cdot \overrightarrow{CP_1} - \overrightarrow{CP_2}$$

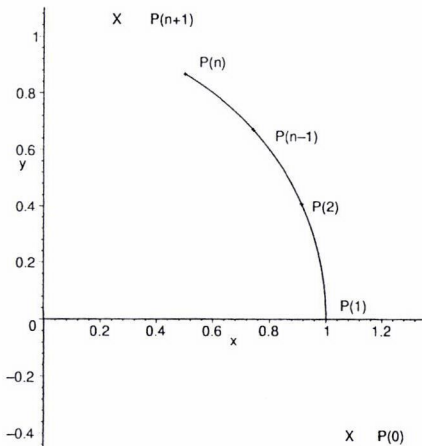
összefüggéssel határozzuk meg.

A  $P_{n+1}$  pontot pedig a

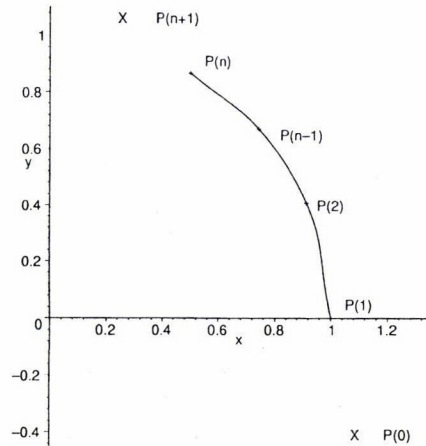
$$p_{n+1} := k + 2 \cdot \cos \delta \cdot \overrightarrow{KP_n} - \overrightarrow{KP_{n-1}}$$

képlettel, ahol  $K$  a  $P_n, P_{n-1}, P_{n-2}$  pontok köré írt kör középpontja,  $\delta$  pedig a  $\overrightarrow{KP_{n-1}}$  és  $\overrightarrow{KP_n}$  vektorok által bezárt szög. (A pontok helyvektorát a megfelelő kisbetűvel jelöltük.)

A 2. ábrán egy körívet és azon adott interpolációs pontokat látunk. A  $P_0$  és  $P_{n+1}$  pontokat a természetes peremfeltételt felhasználva vettük fel. A 3. ábrán jól látható, hogy a  $P_1$  és  $P_2$  pontok között a konvex kombináció után létrejött trigonometrikusan súlyozott spline-görbe görbülete eltér az eredeti görbéétől, mert a  $P_0, P_1$  és  $P_2$  pontokon átmenő egyenes „kiegyenesítette” a görbét. Ugyanez figyelhető meg a  $P_{n-1}$  és  $P_n$  pontok között is.



2. ábra. Köríven választott interpolációs pontok



3. ábra. Adott pontokra illesztett interpolációs görbe

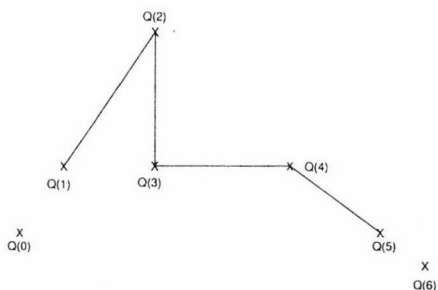


## 2.3.

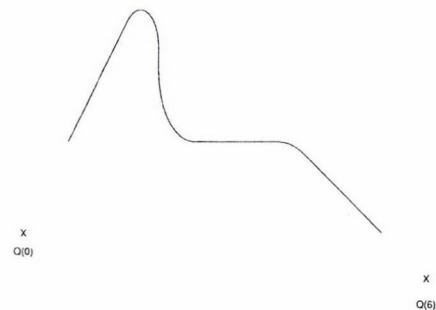
Tekintsünk egy példát a trigonometrikusan súlyozott interpolációs spline-görbe alkalmazására. Egy poligon csúcsait kerekítjük le adott sugarú körívvel. Adott a poligon a csúcspontjaival:  $\{Q_1, Q_2, Q_3, \dots, Q_n\}$  és a lekerekítési körök  $r_i$  ( $i = 2, \dots, n-1$ ) sugara. Az adott töröttvonal mentén felvesszük az interpolációs pontokat, amelyek a lekerekítő körívnek a poligon oldalain lévő érintési pontjai, és azoktól adott, elég kicsi  $\epsilon$  távolságra mindkét irányban további egy-egy pontot a köríven és a poligon oldalán [4, 5, 6].

Legyen  $S := \{P_1, P_2, P_3, \dots, P_k\} \subset \mathbf{E}^3$  ( $k = (n-2)*6+2$ ) az interpolációs pontok sorozata, amely a spline-görbe meghatározó adata. Ebben az esetben a  $Q_1$  és  $Q_n$  pontoknál nincs csúcs, ezeket nem kell lekerekíteni, így a természetes peremfeltételt használhatjuk. A  $P_k$  pontokra alkalmazzuk az előzőekben leírt eljárást, és létrehozuk a  $P_1, P_2, P_3, \dots, P_k$  pontokon átmenő trigonometrikusan súlyozott spline-függvényt. Ez a görbe egy érintőlegesen csatlakozó körív és egyenes szakasz párost négy görbeívvvel helyettesít. Nevezetesen az  $\epsilon$  távolsággal mindkét végén megrövidített körív és egyenes szakasz között két görbület-folytonosan csatlakozó átmeneti görbeívet tartalmaz.

A 4. ábrán a poligont, az 5. ábrán a lekerekítéssel kapott görbét látjuk.



4. ábra. Térbeli poligon

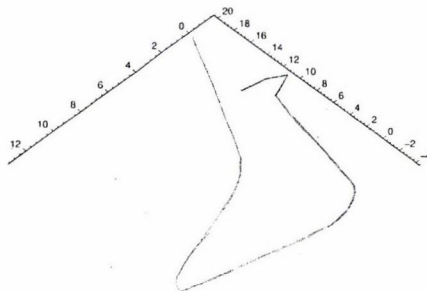


5. ábra. Csúcsoknál lekerekített térbeli poligon

## 3. Vegyes típusú felületfoltokból álló eltolási felület

A 2. pontban előállított trigonometrikusan súlyozott spline legyen a felület vezérgörbéje. Generáló görbeként négy kontrollponttal adott Bézier-görbét választunk, és a vezérgörbe mentén „mozgatjuk” [5] (a 6. ábrán *vastag vonallal a kontrollpoligon látható*).

A görbeívek csatlakozási pontjaiban elhelyezzük a Bézier-görbe kontrollpoligonját. Ennek a térbeli helyzetét szabadon választhatjuk meg. Mivel a vezérgörbe általában nem síkgörbe, a kontrollpoligon alkalmas pozicionálásához a végpontok-



6. ábra. Vezérgörbe és a generáló görbe kontrollpoligonja

hoz lokális koordináta-rendszerként a pontbeli kísérő triédert rendeljük. Kiindulásként tekintjük a vezérgörbe kezdőpontjában megadott kontrollpoligon csúcsait és azok koordinátáit a kezdőpontban felvett lokális koordináta-rendszerben. Feltehetjük, hogy a kísérő triéder a vezérgörbe íveinek kezdő- és végpontjában is létezik.

Legyen  $V_k$  ( $k = 1, \dots, 4$ ) a Bézier-görbe négy alappontja. Egy koordináta-rendszer transzformációval a  $V_1$  pontot a régi koordinátaival a vezérgörbe  $P_i$  interpolációs pontjába, a  $\text{gív}[i](t)$  görbe  $P_i := \text{gív}[i](0)$  pontjában vett kísérő triédere által meghatározott koordináta-rendszerbe helyezzük, legyen ez  $\hat{V}_{i,1}$ . Ugyanez a transzformáció meghatározza a másik három alappontot, és ezáltal a köbös Bézier-görbe kezdőpontja is a görbeív kezdőpontjába kerül. Hasonlóképpen helyezzük el a kontrollpoligont a  $\text{gív}[i](t)$  ív  $P_{i+1} := \text{gív}[i](1)$  végpontjához tartozó kísérő triéder koordináta-rendszerébe is, ez a  $\hat{V}_{i+1,1}$  ponttal kapcsolódik a vezérgörbe  $P_{i+1}$  pontjához. A görbeív közbülső pontjaiban a kontrollpoligon csúcspontjait a két végpontjához rendelt  $\hat{V}_{i,k}$  és  $\hat{V}_{i+1,k}$  ( $k = 1, \dots, 4$ ) pontokból készített konvex kombinációja adja:

$$V_{i,1}(t) = \text{gív}[i](t)$$

$$(2) \quad V_{i,k}(t) = \cos^2\left(t \cdot \frac{\pi}{2}\right) \cdot \hat{V}_{i,k} + \sin^2\left(t \cdot \frac{\pi}{2}\right) \cdot \hat{V}_{i+1,k} + \text{gív}[i](t),$$

$$t \in [0, 1], \quad i = 1, \dots, n-1, \quad k = 2, 3, 4.$$

Ezt minden görbeívre ( $i = 1, \dots, n-1$ ) elvégezzük. Ezáltal a köbös Bézier-görbe kontrollpoligonját végigvisszük a trigonometrikusan súlyozott spline vezérgörbén, majd minden pontban elkészítjük a Bézier-görbét [2]. Az eltolási felületet így vegyes típusú, egyik irányban trigonometrikus, másik irányban pedig polinomiális kétparaméteres vektorfüggvény írja le:

$$(3) \quad \mathbf{R}_i(t, u) = \begin{bmatrix} u^3 & u^2 & u & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} V_{i,1}(t) \\ V_{i,2}(t) \\ V_{i,3}(t) \\ V_{i,4}(t) \end{bmatrix} =$$

$$= [B_{03}(u) \ B_{13}(u) \ B_{23}(u) \ B_{33}(u)] \cdot [V_{i,1}(t) \ V_{i,2}(t) \ V_{i,3}(t) \ V_{i,4}(t)]^T, \\ t, u \in [0, 1], \quad i = 1, \dots, n-1,$$

ahol  $B_{j3}(u)$  ( $j = 0, \dots, 3$ ) súlyfüggvények a Bernstein alappolinomok, amelyek a harmadfokú polinomtér bázisát alkotják.

**3.1. TÉTEL.** *A fent definiált (3) felületfoltok normális-folytonosan kapcsolódnak egymáshoz, vagyis az érintő sík az egyik felületdarabról a másikra való átlépéskor folytonosan változik.*

*Bizonyítás.* Mivel két szomszédos felületdarab egyik irányban egy harmadfokú Bézier-görbe mentén csatlakozik, amely  $C^2$ -folytonos, elegendő megmutatni, hogy a másik paraméter irányában a trigonometrikus paramétervonalak  $GC^1$ , azaz érintő-folytonosak. Ezt a Bézier-kontrollpontok pályagörbéinek érintő-folytonosságával igazoljuk. Tekintsük a  $k$ -adik ( $k = 1, \dots, 4$ ) kontrollpont pályagörbét, és vizsgáljuk a (2) vektorfüggvény  $(i-1)$ -edik és  $i$ -edik ívét:

$$V_{i-1,k}(t) = \cos^2\left(t \cdot \frac{\pi}{2}\right) \cdot \hat{V}_{i-1,k} + \sin^2\left(t \cdot \frac{\pi}{2}\right) \cdot \hat{V}_{i,k} + \text{gív}[i-1](t) \\ t \in [0, 1]$$

$$V_{i,k}(\tau) = \cos^2\left(\tau \cdot \frac{\pi}{2}\right) \cdot \hat{V}_{i,k} + \sin^2\left(\tau \cdot \frac{\pi}{2}\right) \cdot \hat{V}_{i+1,k} + \text{gív}[i](\tau) \\ \tau \in [0, 1]$$

$$i = 1, \dots, n-1, \quad k = 1, \dots, 4.$$

Ezek deriváltjai:

$$\frac{d}{dt}V_{i-1,k}(t) = -\sin(t \cdot \pi) \cdot \frac{\pi}{2} \cdot \hat{V}_{i-1,k} + \sin(t \cdot \pi) \cdot \frac{\pi}{2} \cdot \hat{V}_{i,k} + \frac{d}{dt} \text{gív}[i-1](t) \\ \frac{d}{d\tau}V_{i,k}(\tau) = -\sin(\tau \cdot \pi) \cdot \frac{\pi}{2} \cdot \hat{V}_{i,k} + \sin(\tau \cdot \pi) \cdot \frac{\pi}{2} \cdot \hat{V}_{i+1,k} + \frac{d}{d\tau} \text{gív}[i](\tau) \\ t \in [0, 1], \quad \tau \in [0, 1].$$

A szomszédos görbévek csatlakozási pontjában:

$$\frac{d}{dt}V_{i-1,k}(1) = \frac{d}{dt} \text{gív}[i-1](1) \\ \frac{d}{d\tau}V_{i,k}(0) = \frac{d}{d\tau} \text{gív}[i](0)$$

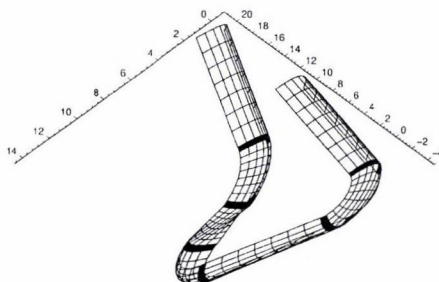
adódik. Ezek az érintő vektorok párhuzamosak és egyirányúak, ahogy azt a 2.1. tételben bizonyítottuk. Tehát a Bézier-kontrollpontok pályagörbéi érintő-folytonosak.  $\square$



Amennyiben a vezérgörbe egy poligon lekerekítésével származtatott, trigonometrikusan súlyozott interpolációs spline, akkor egyenes szakaszokat is tartalmaz. Az egyenes szakaszokhoz viszont mindkét végpontban átmeneti görbeívek csatlakoznak. Végül tehát a görbe minden csatlakozási pontjában létezik a kísérő triéder, ezért a fenti felületgenerálás elvégezhető.

A  $\frac{\partial \mathbf{R}_i(t,u)}{\partial t} \times \frac{\partial \mathbf{R}_i(t,u)}{\partial u} \neq \mathbf{0}$  feltétel tetszőleges geometriai bemenő adatokra általában nem végezhető el, az ilyen szingularitások elkerülése a tervezési folyamatban interaktívan megoldható.

A 7. ábrán egy ilyen vegyes típusú felületfoltokból álló eltolási felületet látunk.



7. ábra. Eltolási felület

#### 4. Két vezérgörbével definiált általánosított eltolási felület

##### 4.1. Két vezérgörbével és egy generáló görbével definiált felületek

Legyen a  $P_{1i}$  ( $i = 1, \dots, n$ ) interpolációs pontokon átmenő trigonometrikusan súlyozott spline-görbe  $g_i(t)$ , és a  $P_{2i}$  ( $i = 1, \dots, n$ ) pontokat interpoláló görbe  $c_i(t)$ , ( $t \in [0, 1]$ ) a két vezérgörbe (11. ábra). A generáló görbe a  $V_k$  ( $k = 1, \dots, 4$ ) kontrollpontokkal adott Bézier-görbe. Az  $i$ -edik ( $i = 1, \dots, n-1$ ) görbeívhez tartozó kontrollpoligont a következőképpen kapjuk:

- a  $V_1$  kontrollpontot a  $g_i[t](t)$  görbe  $g_i[t](0)$  pontjába helyezzük az ezen ponthoz tartozó kísérő triéder koordináta-rendszerében, így  $\hat{V}_{i,1}$  adódik. Ugyanez a transzformáció a  $V_2$  alappontot  $\hat{V}_{i,2}$  pontba viszi szintén a  $g_i[t](0)$  ponthoz tartozó kísérő triéder koordináta-rendszerében.

- A  $V_4$  kontrollpontot a  $c_i[t](t)$  görbe kezdőpontjába helyezzük, a ponthoz tartozó kísérő triéder koordináta-rendszerébe való transzformálással:  $\hat{V}_{i,4}$  adódik, majd ugyanabban a koordináta-rendszerben  $\hat{V}_{i,3}$ -t is kiszámítjuk.

- Így a görbeív elején létrehozuk a  $\hat{V}_{i,k}$  ( $k = 1, \dots, 4$ ) pontokkal adott kontrollpoligont (8. ábra).

- Ezt az ív végén is elvégezzük, s a  $\hat{V}_{i+1,k}$  ( $k = 1, \dots, 4$ ) kontrollpontokat kapjuk.

- A görbeív közbülső pontjaiban a kontrollpoligon a következőképpen adódik:

$$V_{i,1}(t) = \text{gív}[i](t)$$

$$V_{i,2}(t) = \cos^2\left(t \cdot \frac{\pi}{2}\right) \cdot \hat{V}_{i,2} + \sin^2\left(t \cdot \frac{\pi}{2}\right) \cdot \hat{V}_{i+1,2} + \text{gív}[i](t)$$

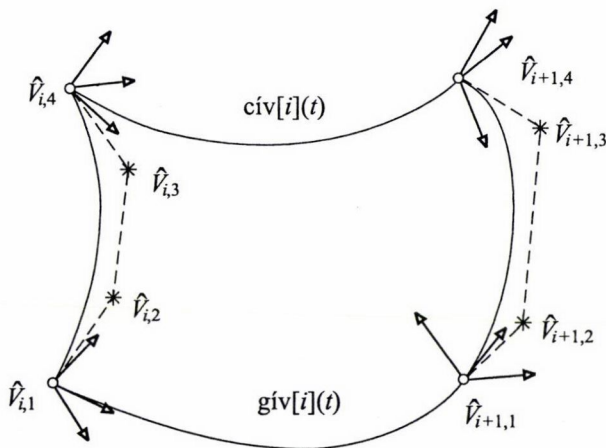
$$V_{i,3}(t) = \cos^2\left(t \cdot \frac{\pi}{2}\right) \cdot \hat{V}_{i,3} + \sin^2\left(t \cdot \frac{\pi}{2}\right) \cdot \hat{V}_{i+1,3} + \text{cív}[i](t)$$

$$V_{i,4}(t) = \text{cív}[i](t)$$

$$t \in [0, 1], \quad i = 1, \dots, n-1.$$

- Ezzel az eljárással a  $\hat{V}_{i,1}$  a  $\text{gív}[i](t)$  görbén, a  $\hat{V}_{i,4}$  pont a  $\text{cív}[i](t)$  görbén „csúszik”.

Az ezen pontokból generált harmadfokú Bézier-görbék halmaza és a két trigonometrikusan súlyozott spline-görbe alkotja a kívánt szabadon formált eltolási felületet [3], amely az egyik irányban trigonometrikus, ( $t \in [0, 1]$ ), a másik irányban köbös polinomiális ( $u \in [0, 1]$ ) felületfoltokból áll. Ezen foltokat is a (3) egyenletben adott  $\mathbf{R}_i(t, u)$  ( $i = 1, \dots, n-1$ ,  $(t, u) \in [0, 1] \times [0, 1]$ ) kétparaméteres vektorfüggvény írja le.



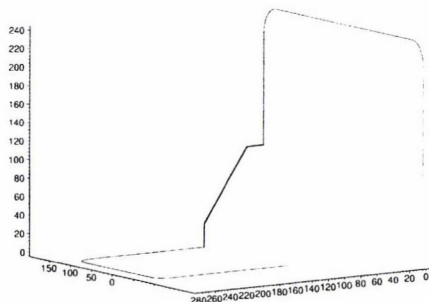
8. ábra. A kontrollpoligon szerkesztése a két vezérgörbe között

4.1. TÉTEL. A Bézier-görbék kontrollpontjainak pályagörbéi érintő-folytonosan csatlakozó görbeívekből állnak.

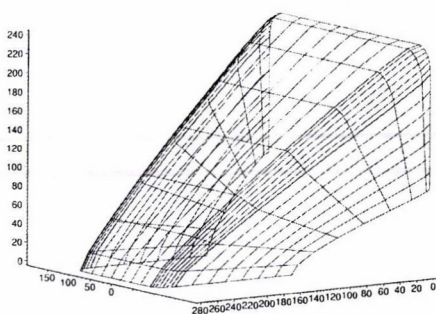
*Bizonyítás.* Ugyanúgy elvégezhető, mint a 3.1. Tétel esetén. □

Így az  $\mathbf{R}_i(t, u)$  ( $i = 1, \dots, n-1$ ,  $(t, u) \in [0, 1] \times [0, 1]$ ) felület  $t$ -paramétervonalai is  $GC^1$ -folytonosak, tehát a felületfoltok normális-folytonosan csatlakoznak

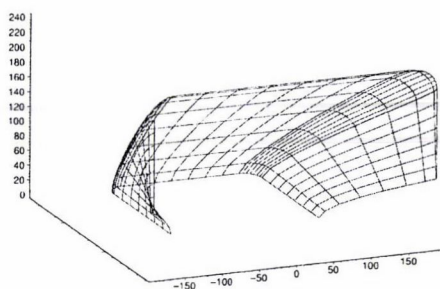
egymáshoz. Ez azt jelenti, hogy a szomszédos felületfoltok felületi normálisai a csatlakozó görbe pontjaiban párhuzamosak és egyirányúak. Ilyenkor valamely felületi görbe mentén haladva az egyik felületdarabról a másikra való átlépéskor az érintő sík folytonosan változik. (9., 10., 11., 12., 13., 14. ábra)



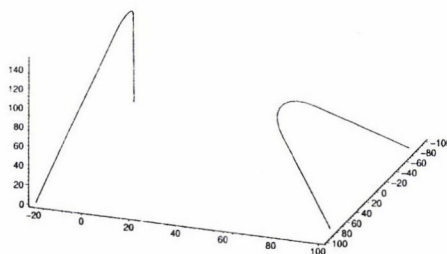
9. ábra. Két vezérgörbe és a generáló görbe kontrollpoligonja



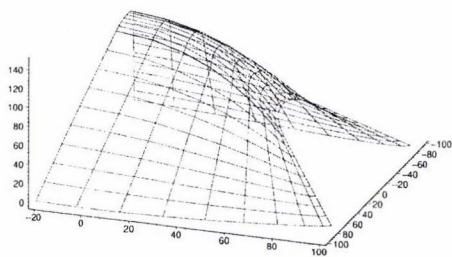
10. ábra. Szabadon formált eltolási felület



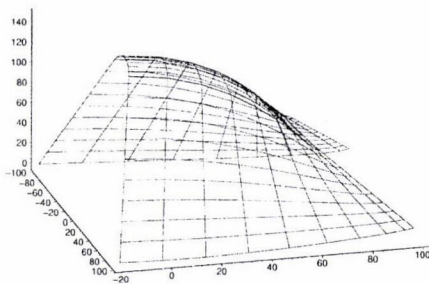
11. ábra. A felület más vetületben



12. ábra. Két vezérgörbe



13. ábra. Szabadon formált eltolási felület a 12. ábra vezérgörbéivel



14. ábra. Ugyanaz a felület más nézetben

#### 4.2. Felületgenerálás alaktartó Bézier-görbével

A két vezérgörbét és a generáló görbét a 4.1. fejezetben leírtak szerint adjuk meg. Egy koordináta-rendszer transzformációval a  $V_1$  kontrollpontot az első, a  $V_4$  kontrollpontot a második vezérgörbe minden görbéjének kezdő- és végpontjához tartozó kísérő triéderének koordináta-rendszerébe helyezzük úgy, hogy a  $\hat{V}_{i,1}$  az első, a  $\hat{V}_{i,4}$  a második vezérgörbén „csússzon”, közben a kontrollpoligon oldalhosszainak aránya az eredetivel megegyezzek. Ez párhuzamos vezérgörbék esetén a kontrollpoligonok hasonlóságát biztosítja. Ehhez minden görbéjén ( $i = 1, \dots, n-1$ ) elején ill. végén meghatározzuk a  $\lambda_1$  ill.  $\lambda_2$  konstans szorzókat ( $\lambda_1, \lambda_2 > 0$ ), amelyek a kontrollpoligon két végpontja közötti aktuális és eredeti távolságainak arányát adják meg. Az  $i$ -edik ( $i = 1, \dots, n-1$ ) görbéjén közbülső pontjaiban a Bézier-görbe kontrollpoligonját a következőképpen definiáljuk:

$$V_{i,1}(t) = \text{gív}[i](t)$$

$$V_{i,2}(t) = \cos^2\left(t \cdot \frac{\pi}{2}\right) \cdot \lambda_1 \cdot \hat{V}_{i,2} + \sin^2\left(t \cdot \frac{\pi}{2}\right) \cdot \lambda_2 \cdot \hat{V}_{i+1,2} + \text{gív}[i](t)$$

$$V_{i,3}(t) = \cos^2\left(t \cdot \frac{\pi}{2}\right) \cdot \lambda_1 \cdot \hat{V}_{i,3} + \sin^2\left(t \cdot \frac{\pi}{2}\right) \cdot \lambda_2 \cdot \hat{V}_{i+1,3} + \text{cív}[i](t)$$

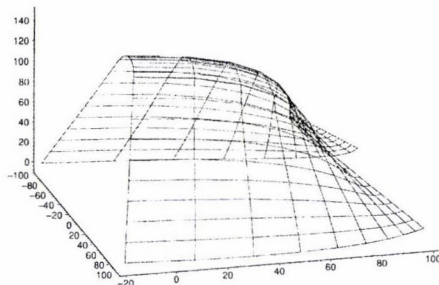
$$V_{i,4}(t) = \text{cív}[i](t)$$

$$t \in [0, 1], \quad i = 1, \dots, n-1.$$

*Megjegyzés.* Ezekkel a kontrollpontokkal generált felületfoltok ugyanúgy  $GC^1$ -folytonosan csatlakoznak egymáshoz mint a 4.1. pontban készített felületfoltok [7].

A 15. ábrán a felületet ugyanazokból az adatokból generáltuk mint a 14. ábrán mutatott felületet, de a kontrollpoligon mozgatása során az oldalhosszak arányát megtartottuk.





15. ábra. Alaktartó transzformációval készült felület

### 4.3. Alakparaméterek meghatározása simasági feltételből

A két vezérgörbével definiált szabadon formált felület létrehozásakor a generáló Bézier-görbe kontrollpoligonja általában torzul. A két görbén ugyanis a megfelelő görbeívek végpontjainak távolsága, vagyis a  $\hat{V}_{i,1}$  és  $\hat{V}_{i,4}$  ( $i = 1, \dots, n-1$ ) kontrollpontok távolsága változik. Másrészt ezekben a pontokban a két görbe kíséző triéderének élei általában nem párhuzamosak, következésképpen az adott  $\hat{V}_{i,1}\hat{V}_{i,2}\hat{V}_{i,3}\hat{V}_{i,4}$  kontrollpoligon  $\hat{V}_{i,1}\hat{V}_{i,2}$  és  $\hat{V}_{i,3}\hat{V}_{i,4}$  éleinek a különböző kíséző triéderek koordináta-rendszereihez való rögzítése a kontrollpoligon torzulását okozza. A 4.2. pontban az oldalhosszak arányát megtartó szorzótényezők bevezetésével – amelyeket a felület alakparamétereinek nevezhetünk – mérsékeljük ezt a torzulást.

Az alakparaméterek megválasztására hatékony módszernek bizonyult a felületek fizikai viselkedését jellemző különböző energiamennyiségek minimalizálása. A gyakorlatban jól bevált a vékony membrán energiáját approximáló

$$F = \int_A \left[ \left( \frac{\partial^2 \mathbf{R}(u, t)}{\partial u^2} \right)^2 + 2 \left( \frac{\partial^2 \mathbf{R}(u, t)}{\partial u \partial t} \right)^2 + \left( \frac{\partial^2 \mathbf{R}(u, t)}{\partial t^2} \right)^2 \right] du dt$$

$$A = [0, 1] \times [0, 1]$$

felületi integrálnak – mint simító függvénynek (fairing function) – a felületmodellezésben való alkalmazása. Ez a függvény a  $\lambda$  alakparaméterekben másodfokú, pozitív definit [1], ezért a minimum helye jól számítható. Az alakparaméterekre tett feltételt, miszerint azok értékét úgy választjuk, hogy az  $F$  függvény értéke minimális legyen, simasági feltételnek (fairness condition) nevezzük.

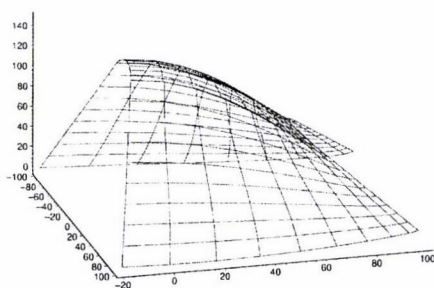
A két vezérgörbével és egy harmadfokú Bézier-görbével definiált szabadon formált felület alakparaméterei a  $\lambda_i$  ( $i = 1, \dots, n-1$ ) szorzótényezők. A simító függvénynek az egész felületre vett integrálja

$$S(\lambda_1, \lambda_2, \dots, \lambda_{n-1}) = \sum_{i=1}^{n-1} F$$

$n - 1$  változós másodfokú polinom. A szélsőérték meghatározása tehát az alábbi lineáris egyenletrendszer megoldására vezet:

$$\frac{\partial S}{\partial \lambda_i} = 0, \quad i = 1, \dots, n - 1.$$

A 16. ábrán látható felületet ugyanazokkal a vezérgörbékkel és Bézier-poligon-nal készítettük, mint a 14. ábrán mutatott felületet, de az alakparamétereket a simasági feltételből számítottuk. Számos példán elvégzett vizsgálatok azt eredményezték, hogy az így generált felületek „feszesebb” alakot mutatnak, mint a 4.1. pontban leírt módszer, ami a  $\lambda_i = 1$  ( $i = 1, \dots, n - 1$ ) esetnek felel meg.



16. ábra. Simasági feltétellel számolt felület

A programot, a számításokat és az ábrákat a MAPLE V számítógépi szimbolikus algebrai program-rendszerrel készítettük.

## Irodalom

- [1] Greiner, G. Variational design and fairing of spline surfaces, *Computer Graphics Forum*, **13** (1994), No. 3, 143–154.
- [2] Nagyné Szilvási Márta, *CAD iskola* (TypoTEX Kft Elektronikus Kiadó, 1991).
- [3] Nagyné Szilvási Márta, *CADKEY Gyakorlókönyv* (Műegyetem Kiadó, 1997).
- [4] Márta Szilvási-Nagy and Teréz P. Vendel, Generating curves and swept surfaces by blended circles, *Computer Aided Geometric Design*, **17** (2000), 197–206.
- [5] M. Szilvasi-Nagy and T. P. Vendel, „Flächenkonstruktion mit trigonometrischen Spline-Funktionen”, *Konstruktive Geometrie Tagung 14.9.–18.9. 1998, Balatonföldvár*.
- [6] Pethőné Vendel Teréz, „Sima görbe generálása körívekből”, *Matematika-, fizika- és informatikaoktatók XXII. Országos Konferenciája* (1998), Budapest.
- [7] Teréz P. Vendel, „Über eine neuartige Translationsfläche”, *Internationale Tagung über Geometrie, Algebra und Analysis, 1.5.–5.5. 1999, Balatonfüred*.

- [8] Wenz H.-J., Interpolation of curve data by blended generalized circles, *Computer Aided Geometric Design*, **13** (1996), 673–680.

(Beérkezett: 1999. október 18.)

PETHŐNÉ VENDEL TERÉZ  
 Pécsi Tudományegyetem  
 Pollack Mihály Műszaki Főiskolai Kar  
 Matematika Tanszék  
 7624 Pécs, Boszorkány U. 2.  
 E-mail: VENDEL@MAT.PMMFK.PTE.HU

## FLÄCHENGENERIERUNG AUS MIT TRIGONOMETRISCHEN BINDEFUNKTIONEN DEFINIERTEN SPLINE-KURVEN

TERÉZ P. VENDEL

Eine räumliche Spline-Kurve wird mit trigonometrischen Binfunktionen definiert, die durch vorgegebene Punkte durchläuft und deren Segmente sich aneinander krümmungstetig anschließen. Dann werden allgemeine Translationsflächen mit einer oder mit zwei Leitkurven definiert, wo die Leitkurven die vorher definierten Spline-Kurven sind. Die erzeugende Kurve ist eine kubische Bézier-Kurve, deshalb sind die Parameterlinien dieser Flächen trigonometrisch in einer Richtung und polynomial in der anderen Richtung. Der Übergang zwischen solchen Flächenstücken ist normalenstetig. Die Seitenlängen des verschobenen Kontrollpolygons der Bézier-Kurve werden einerseits proportional zu den vorgegebenen Seitenlängen gewählt, andererseits werden aus einer Glattheitskriterium berechnet. Die generierten Flächen sind in den mit MAPLE erzeugten Figuren dargestellt.





## ÁLTALÁNOSÍTOTT MULTIFIT TÍPUSÚ MÓDSZEREK II.

DÓSA GYÖRGY

Egyforma párhuzamos gépek ütemezésével foglalkozunk: Hogyan osszunk szét  $n$  munkát  $m < n$  gép között úgy, hogy a teljes átfutási idő a legkisebb legyen: a legkésőbb befejeződő munka a lehető legkorábban fejeződjön be. A Multifit algoritmus [1] egyfajta általánosítást adja [2]. Jelen cikk egy másfajta általánosítási lehetőséggel foglalkozik. A logaritmikus keresés keretét a First Fit Decreasing algoritmus helyett olyan ládapakolási algoritmusokra alkalmazzuk, ahol a ládákat az indexsorrendjük szerint töltjük meg tárgyakkal. Becslést adunk az elméleti hatékonyságra, kísérleti eredményekkel illusztráljuk az új algoritmusok hatékonyságát.

### 1. Bevezetés

Az ütemezéselmélet gyakran vizsgált feladata a párhuzamos gépek ütemezésének problémája, a  $P \parallel C_{\max}$  probléma [6]. Ez a következő: Adott egy feladathalmaz:  $\mathcal{T} = \{T_1, T_2, \dots, T_n\}$ , ahol  $T_i$  ( $i = 1, \dots, n$ ) olyan munkát jelent, amelyet  $m$  gép valamelyikével kell elvégezni. A  $T_i$  munka elvégzésének ideje, vagy röviden hosszúsága  $l(T_i)$ . A munkák elvégzése bármelyik gépen ugyanannyi időbe kerül. Megszakítást nem engedünk meg, vagyis ha egy gép valamelyik munkát elkezd, akkor azt be is kell fejeznie. A munkák egy ütemezésén a  $\mathcal{T}$  halmaz valamely  $\mathcal{P} = \{P_1, P_2, \dots, P_m\}$  partícióját értjük. (Ekkor az  $i$ -edik gép végzi a  $P_i$ -ben lévő feladatokat.) Feltesszük, hogy várakozási idők nincsenek, vagyis amint egy gép befejezi valamelyik általa elvégzendő munkát, és még van olyan, amit neki kell elvégezni, akkor azt rögtön el is kezdi. Az egy-egy gépen elvégzendő munkák sorrendje az ütemezési feladat szempontjából közömbös. A  $\mathcal{P}$  ütemezés teljes átfutási idejét a következőképpen definiáljuk:

$$(1) \quad L(\mathcal{P}) = \max_{1 \leq i \leq m} l(P_i)$$

ahol  $\mathcal{T}$  tetszőleges  $X$  részhalmaza esetén  $l(X) = \sum_{T \in X} l(T)$ , vagyis  $l(P_i)$  az  $i$ -edik gépen a legkésőbb befejeződő munka befejezési idejét jelenti, ha feltesszük, hogy

a gépek a 0 időpontban kezdenek dolgozni. A  $\mathcal{P}^*$  ütemezés optimális, ha teljesül  $L(\mathcal{P}^*) \leq L(\mathcal{P})$  a  $\mathcal{T}$  halmaz tetszőleges  $\mathcal{P}$  ütemezése esetén. Mivel véges sokféleképpen tudjuk a  $\mathcal{T}$  halmaz elemeit  $m$  részre partícionálni, ilyen optimális ütemezés biztosan létezik, esetleg több is lehet. Az  $L(\mathcal{P}^*)$  értéket jelöljük  $C^*$ -gal, ami tehát csak  $\mathcal{T}$ -től és az  $m$  számtól függ.

A feladat téglalappakolási feladatként is megfogalmazható: Itt a munkáknak egy-egy  $l(P_i)$  magasságú, egy egység szélességű téglalap felel meg, ezeket kell egy  $m$  egység szélességű, alul zárt, felül nyitott sávon elhelyezni átfedés nélkül úgy, hogy minimális magasságot használjunk fel a sávból. A téglalapok oldalai a sáv oldalai-val párhuzamosak, és függőlegesen állnak, vagyis a forgatásukat nem engedjük meg. Az optimális ütemezés(ek) megkeresésének feladata NP-teljes [8]. Emiatt a feladat megoldására gyakran korlátozás és szétválasztás típusú módszereket alkalmaznak [7]. Mi itt most gyors, és közel-optimális ütemezéseket adó algoritmusokkal foglalkozunk. Látni fogjuk ugyanis, hogy a heurisztikus módszer is nagyon sok esetben optimális megoldást ad meg. A cikk szerkezete a következő. A 2. paragrafusban korábbi eredményeket vázolunk fel, a harmadik paragrafusban bevezetünk egy új, módosított Multifit-típusú algoritmuscsaládot, amely a korábbi (eredeti) Multifit algoritmus általánosítása. A 4. paragrafusban ennek elméleti hatékonyságára adunk meg felső korlátokat, és végül az 5. paragrafusban teszteredményeket közlünk. Ezek szerint a legtöbb esetben az új algoritmusok jobban működnek, mint az eredeti Multifit algoritmus.

## 2. Korábbi eredmények

R. L. Graham 1966-os [5] cikkéből származik a következő alsó becslés:

$$(2) \quad C^* \geq C_1 := \max \left\{ \frac{l(\mathcal{T})}{m}, \max_{1 \leq i \leq n} l(T_i) \right\},$$

vagyis a teljes átfutási idő legalább akkora, mint a téglalapok összterületének és a sáv szélességének a hányadosa, illetve legalább akkora, mint a leghosszabb téglalap hosszúsága. Ezt az alsó becslést jelöljük  $C_1$ -gyel. Jelöljük egy tetszőleges  $\mathcal{A}$  algoritmus által meghatározott ütemezés átfutási idejét  $L(\mathcal{A})$ -val. (Valójában ez az érték a gépek  $m$  számától, valamint a  $\mathcal{T}$  feladathalmaztól is függ, így az  $L(\mathcal{A}, m, \mathcal{T})$  jelölés lenne teljesen korrekt.) Bevezetjük a következő mennyiséget, amelyet az  $\mathcal{A}$  ütemezési algoritmus elméleti hatékonyságának nevezünk:

$$(3) \quad R_m(\mathcal{A}) = \sup_{\mathcal{T}} \left\{ \frac{L(\mathcal{A})}{C^*} \right\},$$

ahol  $\mathcal{T}$  tetszőleges feladathalmaz.

## 2.1. A MULTIFIT algoritmus

A Multifit algoritmust több szerző javasolta [1]-ben. Itt logaritmikus keresés történik, minden lépésben az FFD (=First Fit Decreasing) ládapakolási algoritmust alkalmazva valamilyen  $C$  láda-mérettel. Az FFD algoritmus a végrehajtási idő csökkenő sorrendje szerint ütemezi a munkákat, vagy a ládapakolás nyelvén kifejezve a ládába a tárgyakat a méretük csökkenő sorrendje szerint helyezi el. A soron következő tárgyat az indexsorrend szerinti legelső olyan ládába teszi, ahova befér. Legyen  $C$  a ládák mérete,  $m$  a ládák száma. Az algoritmus formális leírása a következő:

Az FFD algoritmus

0. Legyen  $P_i = \emptyset$  ( $i = 1 \dots m$ ),  $j = 1$ .
1. Legyen  $i_0 = \min \{i \geq 1, l(P_i) + l(T_j) \leq C\}$ ,  $i_0 > m$  esetén vége.
2. Legyen  $P_{i_0} = P_{i_0} \cup \{T_j\}$ , legyen  $j = j + 1$ .
3.  $j \leq n$  esetén menjünk az 1. lépésre, egyébként vége.

Ha az algoritmus az 1. lépésnél ér véget, akkor túl kicsinek bizonyult a  $C$  ládaméret, nem sikerült az összes tárgyat bepakolni. Legyen  $\text{FFD}[T, C, m] = 1$ , ha az FFD algoritmusnak sikerül bepakolnia a tárgyakat  $m$  ládába, egyébként pedig legyen  $\text{FFD}[T, C, m] = 0$ . Legyen

$$(4) \quad C_2 = \max \left\{ 2 \cdot \frac{l(T)}{m}, \max_{1 \leq i \leq n} l(T_i) \right\}$$

Ekkor [1] szerint  $C < C_1$  esetén  $\text{FFD}[T, C, m] = 0$ ;  $C \geq C_2$  esetén  $\text{FFD}[T, C, m] = 1$  biztosan teljesül. Legyen

$$r_m(\text{FFD}) = \inf \{ r \mid \text{Minden } T\text{-re } \text{FFD}[T, r \cdot C^*, m] = 1 \}.$$

Ekkor érvényes a következő

1. TÉTEL. *Tetszőleges*  $T$  és  $r \geq r_m$  esetén  $\text{FFD}[T, r \cdot C^*, m] = 1$ . □

Ez azt jelenti, hogy ha a ládaméret legalább  $r_m \cdot C^*$ , akkor az algoritmus az összes tárgyat elhelyezi a ládában. A Multifit algoritmus végrehajtásához először megállapítunk egy  $C_L$  alsó-, és egy  $C_U$  felső korlátot a teljes átfutási időre, az előbbiek szerint a  $C_L = C_1$ ,  $C_U = C_2$  választás megfelelő. Ezután a  $[C_L, C_U]$  intervallum  $C$  felezőpontját választjuk a láda méretének, és megpróbáljuk az FFD algoritmus szerint a tárgyakat elhelyezni a ládában. Ha sikerül, vagyis mindegyik tárgy bekerül valamelyik ládába, akkor  $C_U$  szerepét  $C$  veszi át, ha nem sikerül, akkor pedig  $C_L$ -t helyettesítjük  $C$ -vel, és ezt a lépést (egy előre meghatározott  $K$  számszor) iteráljuk.

### A Multifit $[K]$ algoritmus

1. Legyenek  $C_L, C_U$  az átfutási idő alsó, illetve felső becslései, legyen  $i := 1$ .
2.  $i > K$  esetén vége, egyébként  $C := (C_U + C_L)/2$
3. Ha  $\text{FFD}[\mathcal{T}, C, m] = 1$ , akkor  $C_U := C$ ,
4. Ha  $\text{FFD}[\mathcal{T}, C, m] = 0$ , akkor  $C_L := C$ ,
5.  $i := i + 1$ . Menjünk a 2. lépésre.

A Multifit algoritmus elméleti hatékonyságáról [1]-ben a következő becslés szerepel:  $R_m(\text{Multifit}[K]) \leq r_m(\text{FFD}) + \left(\frac{1}{2}\right)^K$ , valamint  $r_m(\text{FFD}) \leq 1.22$ . [3] az utóbbi becslést 1.2-re javította, az algoritmust javítva  $\frac{72}{61}$  pontos felső becslést adott [4].

Megjegyezzük, hogy [2] az előbbi Multifit algoritmus egyfajta általánosítását adja meg, megtartva azt a tulajdonságot, hogy a belső ládapakolási algoritmus a tárgyakat monoton csökkenő méret szerint helyezi el, és a Multifit valamint a Graham-féle LPT algoritmus közös általánosításához jut el. Jelen cikk egy másfajta általánosítási lehetőséggel foglalkozik, most a gépeket egymás után töltjük fel munkákkal balról jobbra haladva, vagyis az indexeik növekvő sorrendjében. A hagyományos Multifit algoritmus tárgyalható ilyen módon is, és ez az átfogalmazás teszi lehetővé azt a fajta általánosítást, amely a továbbiakban következik.

**2. TÉTEL.** *Tekintsük a következő algoritmust: Csökkenő méret szerint haladva tegyük bele az első ládába azokat a tárgyakat, amelyek beférnek még a már behelyezett tárgyak mellé. Ha már nem fér több az első ládába, akkor kezdjük el hasonló módon feltölteni a következő ládát a megmaradt tárgyakkal, és így haladjunk tovább. Ekkor ez az algoritmus az FFD algoritmussal megegyező ládapakolást eredményez.*  $\square$

### 3. Az általánosított Multifit algoritmus

Az FFD algoritmust cseréljük ki egy F algoritmussal, amelyre általános keretet adunk meg. Az algoritmust további feltételek bevezetésével egyre inkább specializálni fogjuk, mindig éppen csak annyira, ahogy a soron következő tételek a szigorítást megkívánják. A ládákat egymás után töltjük fel tárgyakkal az indexeik növekvő sorrendjében haladva. Először kiválasztjuk azokat a részhalmazait a tárgyak  $\mathcal{T}$  halmazának, amely tárgyak még nem lettek elhelyezve, és egyszerre beférnek az aktuális ládába. Ezután ezek közül egy  $R$  értékkel függvény segítségével kiválasztunk egy tárgyhalmazt, és az elemeit a ládába tesszük. Ezt a két lépést addig iteráljuk, amíg van még olyan tárgy, amelyik az aktuális ládába befér, (vagyis a beférő tárgyak halmaza nem üres). Ha már az aktuális ládába nem fér be több tárgy a megmaradtak közül, akkor kezdjük el a következő láda feltöltését. Az algoritmus akkor ér véget, ha elfogynak a tárgyak, tehát mindegyiket sikerült bele-

tenni valamelyik ládába; vagy elfogynak a ládák, és valamely tárgyak kimaradtak a ládákból. Formálisan:

Az F algoritmus

1. Legyen  $C$  a ládák mérete,  $m$  a ládák száma,  $P_i = \emptyset$  ( $i = 1 \dots m$ ),  $\gamma \in N$  egy rögzített szám.
2. Legyen  $i := 1$ ,  $S = T$
3. Legyen  $H = \{X \mid X \subseteq S, l(P_i) + l(X) \leq C\}$ , vagyis  $H$  azon tárgyhalmazok halmaza, amelyek egyszerre beférnek az aktuális ládába.
- 4.1  $H = \emptyset$  esetén legyen  $i = i + 1$ , menjünk a 6. lépésre.
- 4.2  $H \neq \emptyset$  esetén legyen  $X_0 = \arg \max \{R(X) \mid X \in H, |X| \leq \gamma\}$
5. legyen  $P_i = P_i \cup X_0$ ,  $S = S \setminus X_0$ , menjünk a 3. lépésre.
6.  $i \leq m$  esetén menjünk a 3. lépésre, egyébként vége.

Ha a 4.2 pontbeli maximum nem lenne egyértelmű, akkor a maximális értéket adó  $X$  halmazok közül valamilyen szabály segítségével választunk ki egyet.

*Példa.* Legyen  $\gamma = 1$  és  $R(X) = l(X)$ . Ekkor a 2. Tétel szerint éppen az FFD algoritmust kapjuk, a tárgyak a legelső olyan ládába kerülnek, ahova beférnek. Ezek szerint az általános F algoritmusunk az FFD algoritmus általánosítása.

*Definíció.* Legyen  $F$  egy tetszőleges ládapakolási algoritmus. Ha az  $F$  algoritmusnak sikerül bepakolnia a tárgyakat  $m$  számú,  $C$  méretű ládába, akkor legyen  $F[T, C, m] = 1$ , egyébként legyen  $F[T, C, m] = 0$ .

3. TÉTEL. Tetszőleges  $T$  feladathalmaz,  $R$  értékelő függvény és  $C < C^*$  esetén teljesül  $F[T, C, m] = 0$ . □

KÖVETKEZMÉNY. Tetszőleges  $T$  feladathalmaz,  $R$  értékelő függvény és  $C < C_1$  esetén  $F[T, C, m] = 0$ , ahol  $C_1$  a (2)-beli alsó becslés.

*Bizonyítás.* Ez abból következik, hogy  $C_1 \leq C^*$ . □

4. TÉTEL. Tetszőleges  $T$  feladathalmaz,  $R$  értékelő függvény és  $C \geq 2 \cdot C_1$  esetén  $F[T, C, m] = 1$ .

*Bizonyítás.* Tegyük fel, hogy  $C \geq 2 \cdot C_1$ , és  $F[T, C, m] = 0$ . Ez azt jelenti, hogy az  $F$  algoritmus elvégzése után maradt olyan tárgy, amely a ládákban nem lett elhelyezve. Ez a tárgy már semelyik ládába nem fér be. Feltehető, hogy már csak egyetlen tárgy,  $T_n$  maradt elhelyezetlenül, különben a többit elhagyjuk a  $T$  feladathalmazból, és a maradékra még mindig teljesül a kezdeti feltétel, hiszen bizonyos tárgyak elhagyása által a  $C_1$  alsó becslés értéke csak csökkenhet,  $R$  pedig ugyanazokat a téglalapokat választja ki. Ekkor tehát

$$l(P_i) + l(T_n) > C \quad i = 1 \dots m,$$

ahol  $T_n$  az utolsó, el nem helyezett tárgy. Az egyenlőtlenségeket összegezve:

$$\sum_{i=1}^m l(P_i) + m \cdot l(T_n) > m \cdot C.$$

Ha a bal oldalhoz  $l(T_n)$ -t hozzáadunk, (ami nemnegatív), az egyenlőtlenséget  $m$ -mel osztva a következőt kapjuk:

$$(5) \quad \frac{l(T)}{m} + l(T_n) > C \geq 2 \cdot C_1.$$

Másrészt  $C_1 \geq \frac{l(T)}{m}$  és  $C_1 \geq \max_{1 \leq i \leq n} l(T_i) \geq l(T_n)$  miatt

$$(6) \quad 2 \cdot C_1 \geq \frac{l(T)}{m} + l(T_n),$$

ami viszont ellentmond (5)-nek. □

*Megjegyzés.* Az eredeti Multifit algoritmus esetében a tételben szereplő  $2 \cdot C_1$  konstans szerepét a (4) képletben álló  $C_2 \leq 2 \cdot C_1$  játszotta, de a bizonyítás a mostani esetben  $C_2$  esetén nem működik. Másrészt a 2-es szorzó a lehető legkisebb: Álljon a  $T$  halmaz  $m$  számú  $m - 1$  hosszúságú, valamint egy darab  $m$  hosszúságú téglalapról. Ekkor  $C_1 = m$ ,  $C^* = 2 \cdot (m - 1)$ . Ezért  $C < \frac{2 \cdot (m-1)}{m} \cdot C_1$  esetén  $F[T, C, m] = 0$  lesz, vagyis  $C_1$  együtthatója nem lehet kisebb 2-nél. Ezek alapján most is meg tudunk adni alsó és felső korlátot a  $C$  átfutási időre, amelynél kisebb, illetve amelynél nagyobb értékek esetén biztos, hogy  $F[T, C, m] = 0$ , illetve  $F[T, C, m] = 1$ . Az algoritmusunk általános kerete így a következő lesz:

A GMF[K] (általánosított Multifit) algoritmus

1. Legyenek  $C_L \leq C^* \leq C_U$  az átfutási idő alsó, illetve felső becslései, valamint legyen  $i := 1$ .
2.  $i > K$  esetén vége, egyébként  $C := (C_U + C_L)/2$
3. Ha  $F[T, C, m] = 1$ , akkor  $C_U := C$ ,
4. Ha  $F[T, C, m] = 0$ , akkor  $C_L := C$ ,
5.  $i := i + 1$ . Menjünk a 2. lépésre.

*Definíció.* Tetszőleges  $F$  ládapakolási algoritmus esetén definiáljuk a következő mennyiséget:

$$r_m(F) = \inf \{ r \mid \text{Minden } T\text{-re } F[T, r \cdot C^*, m] = 1 \}.$$

Az  $r_m$  szám tehát az a legkisebb szorzó, amennyivel megszorozva a  $C^*$  értéket a kapott méretű  $m$  számú ládába az  $F$  algoritmus biztosan be tudja pakolni a téglalapokat. Könnyen látszik, hogy  $r_m$  definíciója értelmes, hiszen  $C^* \geq \frac{1}{m} l(T)$ , ugyanis

ha a gépekre egyenletesen sikerülne elhelyezni a munkákat, akkor kapnánk ezt az értéket, ekkor viszont  $m \cdot C^* \geq l(T)$ , vagyis  $r = m$  választással minden munka az első gépre kerül, vagyis nem üres számhalmaz infimumát vesszük az előbbi definícióban.

5. TÉTEL. Tetszőleges  $\mathcal{R}$  kiválasztási függvény esetén  $1 \leq r_m(F) \leq 2$ .

*Bizonyítás.* Az előző tétel szerint  $F[T, C, m] = 1$ , ha  $C \geq 2 \cdot C_1$ , másrészt  $C_1 \leq C^*$ , ezért tetszőleges  $C \geq 2 \cdot C^*$  esetén  $F[T, C, m] = 1$ .  $\square$

Ezek után a Multifit-típusú algoritmuscsalád általánosságát fokozatosan korlátozzuk, mindig csak annyira, hogy a következő tétel a lehető legáltalánosabb keretek között teljesüljön.

F1 FELTEVÉS. Legyen  $X_1, X_2 \in H$  az általános Multifit algoritmus 4.2 lépésénél. Ha  $|X_1| \leq \gamma$ ,  $|X_2| \leq \gamma$ , valamint  $l(X_1) < l(X_2)$ , akkor  $X_0 \neq X_1$ .

F2 FELTEVÉS. Legyen  $\varepsilon < \frac{\delta}{n \cdot \gamma}$ , ahol

$$\delta = \min \{ |l(X_1) - l(X_2)| > 0 : X_1 \subseteq T, X_2 \subseteq T \}.$$

Az  $X$  téglalaphalmaz álljon valahány  $\varepsilon$  méretű téglalapból. Legyen  $X_1 \in H$ ,  $|X_1| \leq \gamma$  az algoritmus 4.2 lépésénél. Ekkor ha a  $H$  halmazból a kiválasztási szabály szerint kiválasztott téglalaphalmaznak az  $X_1$  nem részhalmaza, akkor ha a  $H$  halmazt az  $X$  halmazzal kibővítjük, akkor a kibővített halmazból kiválasztott téglalaphalmaznak sem lesz az  $X_1$  halmaz részhalmaza.

*Megjegyzés.* Az előbbiek szerint az  $\mathcal{R}$  értékelő függvény (és az egyértelműséget biztosító pótlólagos kiválasztási szabály) szívesebben választ nagyobb összméretű halmazt. A második feltétel azt jelenti, hogy van olyan kicsi  $\varepsilon$ , amelynél kisebb, egyforma méretű téglalapokkal bővítve a téglalapok halmazát az algoritmus lényegében ugyanúgy működik: az eredeti téglalapokból ugyanazokat helyezi el az egyes ládába, mint az előbb. Az előbbi két tulajdonság az FFD algoritmus esetében triviálisan teljesül. Ez azt jelenti, hogy ha az előbbi feltételek teljesülnek, akkor ugyan korlátoztuk az F algoritmus általánosságát, de az még mindig általánosítása az FFD algoritmusnak.

6. TÉTEL. Tetszőleges  $T$  és  $r \geq r_m$  esetén  $F[T, r \cdot C^*, m] = 1$ , ha feltesszük, hogy a  $T$ -beli tárgyak méretei összemérhetőek, valamint teljesül az F1 és F2 tulajdonság.

*Bizonyítás.* Először belátjuk, hogy  $F[T, r_m \cdot C^*, m] = 1$ . Tegyük fel ezzel el-lentétben, hogy az F algoritmus nem képes a tárgyakat a  $C = r_m \cdot C^*$  méretű,  $m$  számú ládába bepakolni, vagyis legalább egy tárgy kimarad a ládákból. Legyen  $C_\alpha = \min \{ l(X), l(X) > C, X \subseteq T \}$ . (Emlékeztetünk arra, hogy tetszőleges  $X \subseteq T$  részhalmaz esetén  $l(X) = \sum_{T \in X} l(T)$  az  $X$ -beli tárgyak méreteinek összegét jelenti.) Ekkor tetszőleges  $C'$  esetén, ahol  $C \leq C' < C_\alpha$ , a tárgyakat az F algoritmus a  $C'$  méretű ládába sem képes bepakolni, mert minden ládába pontosan

ugyanazok a tárgyak kerülnek, mint az előbb, hiszen a ládába beférő téglalapalmazatok nem változtak; ez pedig ellenmond  $r_m$  értelmezésének. Most lássuk be, hogy  $F[T, r \cdot C^*, m] = 1$  teljesül tetszőleges  $r > r_m$  esetén. Először feltesszük, hogy  $\frac{r}{r_m}$  racionális szám. Tegyük fel tehát, hogy  $F[T, r \cdot C^*, m] = 0$ , valamely  $T$  feladathalmaz esetén, legyen  $r = \alpha \cdot r_m$ ,  $\alpha > 1$ ,  $\alpha \in Q$ . Tekintsünk egy tetszőleges optimális pakolást  $m$  darab,  $C^*$  méretű ládába. Növeljük meg a ládák méretét  $\alpha \cdot C^*$ -ra. Töltsük ki az összes ládában a megmaradt helyeket az  $F_2$  feltételben szereplő  $\varepsilon$ -nál kisebb egyforma méretű tárgyakkal úgy, hogy valamennyi láda teljesen legyen megtöltve. Az új tárgyak mérete legyen az összes korábbi tárgy méretének, valamint  $\alpha \cdot C^*$ -nak is osztója. (Ez elérhető, mert  $\alpha$  racionális.) A kapott feladathalmazt jelöljük  $\bar{T}$ -vel. Ekkor  $\bar{C}^* = \alpha \cdot C^*$ , ahol  $\bar{C}^*$  az új feladathalmaz esetén a teljes átfutási idő optimális értéke. Hajtsuk végre az  $F$  algoritmust az  $r \cdot C^*$  ládamérettel és  $\bar{T}$  feladathalmazzal.  $F_1$  és  $F_2$  teljesülése folytán minden ládába ugyanazok a tárgyak kerülnek mint az előbb, esetleg még néhány az új,  $\varepsilon$  méretű tárgyból. Ezért a következő ládába teendő tárgyak választásánál lényegében ugyanabból a halmazból választunk, nevezetesen az előbbi maradék tárgyak, és még a kicsi fajtából megmaradt tárgyak közül. Mivel az  $F$  algoritmus előbb a régebbi tárgyakat nem tudta mindet elhelyezni, hiszen feltettük, hogy  $F[T, r \cdot C^*, m] = 0$ , ezért most sem tudja, így  $F[\bar{T}, r \cdot C^*, m] = 0$  következik. Mivel  $r \cdot C^* = \alpha \cdot r_m \cdot C^* = r_m \cdot \bar{C}^*$ , így kapjuk:  $F[\bar{T}, r_m \cdot \bar{C}^*, m] = 0$ , ez pedig ellentmond az előző résznek. Ezután tetszőleges  $r > r_m$ -re a bizonyítás könnyen adódik, hiszen az  $F[T, C, m]$  kifejezés értéke a bizonyítás első részében leírtak alapján valamilyen  $[C_\alpha, C_\beta]$  alulról zárt, felülről nyílt intervallumokon ugyanaz.  $\square$

7. TÉTEL.  $R_m(GMF[K]) \leq r_m(F) + \left(\frac{1}{2}\right)^K$ .

*Bizonyítás.* Megegyezik az eredeti esetben történő bizonyítással [1].  $\square$

KÖVETKEZMÉNY.  $R_m(GMF[K]) \leq 2 + \left(\frac{1}{2}\right)^K$

*Bizonyítás.* A 4. Tétel miatt tetszőleges  $T$  és  $C \geq 2C_1$  esetén  $F[T, C, m] = 1$ , ezért  $F[T, 2 \cdot C^*, m] = 1$ , így  $r_m(F) \leq 2$ .

#### 4. Az $F_\gamma$ és $F_\infty$ algoritmus-család

Tovább korlátozzuk az algoritmuscsalád általánosságát: bevezetjük az  $F_\gamma$  és  $F_\infty$  algoritmusokat, amelyek még mindig az FFD algoritmus általánosításai. Az  $F_\gamma$  algoritmus 4.2. -edik lépésében legyen  $R(X) = l(X)$ , vagyis válasszunk ki legfeljebb  $\gamma$  darab tárgyat, amelyek beférnek az aktuális ládába, és az összméretük maximális. Ehhez hasonlóan definiáljuk az  $F_\infty$  algoritmust: Helyezzünk el az aktuális ládába valahány tárgyat úgy, hogy az összméretük legyen maximális. Ez utóbbi algoritmus nyilván azonos tetszőleges olyan  $F_\gamma$  algoritmussal, ahol például  $\gamma$  nagyobb, mint a téglalapok száma, tehát  $F_\infty$  is egy speciális  $F_\gamma$  algoritmus. Ha a maximális összméret többféleképpen is fölvetetik, akkor az elhelyezendő tégl-



laphalmazt ezek közül a következő szabály segítségével válasszuk ki: A  $T$  halmaz elemeit rakjuk a méretük szerinti csökkenő sorrendbe. Azonos összméret esetén válasszunk minimális számú tárgyat. Ha még mindig többféle tárgyhalmaz marad, akkor ezen tárgyhalmazok elemeinek méreteiből alkotott vektorok közül válasszuk a lexikografikusan legnagyobbat.

*Megjegyzés.* Belátható, hogy az  $F_1$  és  $F_2$  feltételek teljesülnek, ezért az előző tétel állítása biztosan teljesülni fog ezen algoritmusok esetén is. A későbbi hatékonyságbecsléseknél az egyértelműséget biztosító megszorításokat nem használjuk fel a bizonyítás során, csak a numerikus vizsgálatoknál alkalmazzuk majd az előbbi szabályt. Egy másik lehetőség, az egyértelműséget biztosító szabály lehetne az is, hogy azonos összméret esetén válasszunk olyan halmazt, amelyikben a „kisebb tárgyak száma kisebb”, ezáltal azt remélve hogy jobban járunk, ha előbb a nagyobb tárgyakat helyezzük el, mert ezeket később nehezebben tudnánk elhelyezni, mint a kisebbeket. Ehhez hasonló szabályok alkalmazása és vizsgálata későbbi vizsgálatok tárgya marad. Ezek után az  $r_m(F)$  konstansra szeretnénk felső becslést kapni, hiszen így egyben a 7. Tétel alapján az  $R_m(GMF[K])$  expanziós faktort is megtudtuk becsülni.

#### 4.1. Hatékonysági becslések

Az  $F_\gamma$  és  $F_\infty$  algoritmusokra vonatkozó felső, illetve pontos hatékonysági becsléseket adunk meg. Ehhez némi előkészítésre van szükségünk.

8. LEMMA. Tegyük fel, hogy az  $F$  algoritmust  $m$  ládára alkalmazva tárgyaknak a következő elhelyezését kapjuk:  $P = \langle P_1, \dots, P_m \rangle$ . Ekkor az  $F$  algoritmust  $m - 1$  ládára alkalmazva a  $T' = T \setminus P_\alpha$ -beli tárgyaknak a következő elhelyezését kapjuk:  $P' = \langle P_1, \dots, P_{\alpha-1}, P_{\alpha+1}, \dots, P_m \rangle$  tetszőleges  $1 \leq \alpha \leq m$  egész szám esetén, vagyis a többi ládába ugyanazok a tárgyak kerülnek, mint az előbb.  $\square$

*Definíció.* Legyen  $p, q$  tetszőleges pozitív egész szám, ahol  $p \geq q$ . Egy  $(p/q)$  ellenpéldán olyan  $(T, M)$  rendezett párt értünk, (ahol  $T$  egy feladathalmaz,  $M$  pedig egy pozitív szám, a ládák száma), amelyekre teljesül  $F[T, p, M] = 0$ , valamint  $C^*[T, M] \leq q$ , vagyis az  $F$  algoritmus nem képes a tárgyakat  $M$  számú  $p$  méretű ládába elhelyezni, de van legfeljebb  $q$  átfutási idővel rendelkező ütemezés.

9. LEMMA. Legyen  $(T, M)$  egy tetszőleges  $(p/q)$  ellenpélda. Ekkor tetszőleges  $\beta > 0$  esetén a  $(T, M)$  pár egyben  $(\beta \cdot p/\beta \cdot q)$  ellenpélda is.  $\square$

*Definíció.* Minimális  $(p/q)$  ellenpéldán olyan  $(T, M)$  rendezett párt értünk, amely a következő értelemben minimális:

- a, A  $(T, M)$  pár  $(p/q)$  ellenpélda
- b, tetszőleges a-t kielégítő  $T'$  esetén  $|T'| \geq |T|$
- c, tetszőleges  $1 \leq m < M$  esetén  $r_m \leq p/q$ .

10. LEMMA. Legyen  $(T, M)$  minimális  $(p/q)$  példa. Ekkor van olyan  $T \in T$  amelyre  $F[T \setminus \{T\}, p, M] = 1$ , vagyis egyetlen tárgy marad ki a ládák közül.  $\square$

4.2.  $m = 2$  esete

Ha a ládák száma kettő, akkor könnyen adhatunk pontos becslést az algoritmusaink elméleti hatékonyságára.

11. TÉTEL.  $r_2(F_\gamma) = \frac{2\gamma+4}{2\gamma+3}$ .

*Megjegyzés.*  $\gamma = 2$  esetén az előbbi becslés értéke  $\frac{8}{7}$ . Ez pontosan annyi, mint az FFD algoritmusra vonatkozó pontos becslés értéke két láda esetén. Nagyobb  $\gamma$  esetén azonban egyre jobb értékeket kapunk, vagyis két gép esetén minél nagyobb a  $\gamma$  paraméter, annál jobb az algoritmusunk elméleti hatékonysága, és  $\gamma$  növelése esetén az elméleti hatékonyság 1-hez tart.

*Bizonyítás.* Álljon a  $T$  halmaz  $2\gamma + 2$  számú téglalaphból, ezek közül kettőnek a mérete 3 egység, a többi hossza 2 egység. Ekkor ezek éppen beleférnek kettő darab,  $2\gamma + 3$  méretű ládába: mindkét ládába egy nagyobb téglalap, és  $\gamma$  darab kisebb kerül. Ekkor mindegyik ládában  $\gamma + 1$  darab téglalap van. Ha most legfeljebb  $\gamma$  darab maximális összméretű téglalapot akarunk beletenni egy  $2\gamma + 4 - \varepsilon$  méretű ládába (ahol  $0 < \varepsilon < 1$ ), akkor pontosan  $\gamma$  darabot helyezünk el, a két darab 3 hosszúságút, és  $\gamma - 2$  darab 2 hosszúságút. Ezek együttes mérete  $2\gamma + 2$ . Ezek mellé már nem fér el téglalap, és a többi pedig nem fér be a másik ládába. Ezért  $r_2(F_\gamma) \geq \frac{2\gamma+4}{2\gamma+3}$ . Most belátjuk, hogy  $r_2(F_\gamma) \leq \frac{2\gamma+4}{2\gamma+3}$ . Jelöljük a ládákat  $P_1$  és  $P_2$ -vel, az  $F_\gamma$  algoritmus végrehajtása során a  $P_1$ -be kerülő tárgyak halmazát  $T_1$ -gyel, a  $P_1$ -ből kimaradó tárgyak halmazát (amelyek a  $P_2$ -be kerülő tárgyak, valamint a mindkét ládából kimaradó tárgy) pedig jelöljük  $T_2$ -vel. A Tétel bizonyításának további részéhez szükségünk lesz a következő állításra:

12. SEGÉDTÉTEL.  $|T_2| \leq \gamma + 1$ .

*Bizonyítás.* Tegyük fel, hogy legalább  $\gamma + 2$  darab téglalap kimarad  $P_1$ -ből, ezek együtt nem férnek el  $P_2$ -ben. Az összterületük legyen  $2\gamma + 4 + \varepsilon$ , ahol  $\varepsilon > 0$ . Ekkor  $l(P_1) \leq 2\gamma + 2 - \varepsilon$ . Mivel a  $P_1$ -ből kimaradó tárgyak egyike sem fér már be  $P_1$ -be, mindegyik  $T_2$ -beli tárgy mérete nagyobb, mint  $2 + \varepsilon$ . Ezért a  $T_2$ -beliek összmérete nagyobb, mint  $(\gamma + 2)(2 + \varepsilon) = 2\gamma + 4 + (\gamma + 2)\varepsilon$ , ami ellentmondás.  $\square$

A tétel bizonyítása ezután a következő: Legyen  $T'$  az a legfeljebb  $\gamma$  darab téglalap, amelyet az  $F_\gamma$  algoritmus először elhelyez az első ládában. Legyen  $T^* = T' \cup T_2$ . Az előbbi Lemma alapján  $|T^*| \leq 2\gamma + 1$ . A  $T^*$ -beli tárgyak elhelyezhetők két darab  $2\gamma + 3$  méretű ládában, ekkor ezek között van olyan láda, amelyikben legfeljebb  $\gamma$  darab tárgy van, ezen tárgyak összmérete legyen  $C$ . Ezek szerint ha az egyik ládába a  $T^*$ -beli tárgyak közül legalább  $C$  területnyi kerül, akkor a másik ládába befér a többi  $T^*$ -beli tárgy. Amikor  $F_\gamma$  elhelyezte a  $T'$ -beli tárgyakat az első ládába, itt az összterület ezáltal már legalább  $C$  volt. Ezért a  $T_2$ -beliek elférnek a másik ládában, ami ellentmondás.  $\square$

13. TÉTEL.  $r_2(F_\infty) = 1$ .

*Bizonyítás.* Legyen  $(T, M)$  tetszőleges  $(p/q)$  ellenpélda. Tekintsünk egy tetszőleges pakolást két  $q$  méretű ládába, az egyik ládába tett tárgyak összmérete legyen

C. Az  $F_\infty$  algoritmus legalább  $C$  összméretű tárgyat pakol az első  $p$  méretű ládába, ezért a többi tárgy befér a másik ládába. Ellentmondásra jutottunk, ez azt jelenti, hogy  $(p/q)$  ellenpélda nem létezik.  $\square$

*Megjegyzés.* Az előbbi bizonyítási technika több láda esetén sajnos nem működik, ezt mutatja az alábbi példa. Álljon  $T$  a következő tárgyakból:  $T = \{7, 7, 5, 3, 2, 2\}$ , legyen  $P_1 = \{5, 3\}$ ,  $P_2 = \{7, 2\}$ ,  $P_3 = \{7, 2\}$ , a ládák mérete 9. Ha az első ládába nagyobb összméretű tárgyakat teszünk a következőképpen:  $P'_1 = \{5, 2, 2\}$ , akkor az első ládából kimaradó három tárgy már nem fér el a másik két ládában. Vagyis, ha nagyobb összméretű tárgy kerül az első, vagy előbbi ládába, ez nem garantálja azt, hogy a többi tárgy még inkább befér a hátralévő ládába.

### 4.3. $m > 2$ esete

Ha a ládák száma legalább három, újra szükségünk lesz némi előkészítésre.

*Definíció.* Azt mondjuk, hogy  $X, Y \subseteq T$  esetén  $X$  dominálja  $Y$ -t, ha létezik olyan  $f : Y \rightarrow X$  injektív leképezés, amelyre  $l(f(y)) \geq l(y)$  tetszőleges  $y \in Y$  esetén.

14. LEMMA. Legyen  $(T, M)$  minimális  $(p/q)$  ellenpélda. Legyen  $P = \langle P_1, \dots, P_M \rangle$  az  $F$  algoritmus által történő elhelyezése a tárgyaknak, ahol egy tárgy kimarad a ládából, ezt az egyetlen tárgyat tegyük bele a  $P_{M+1}$  ládába,  $P^* = \langle P_1^*, \dots, P_M^* \rangle$  pedig legyen egy tetszőleges optimális pakolás. Ekkor a  $P_i$  halmaz nem dominálja a  $P_j^*$  halmazt, ahol  $i \in \{1, \dots, M+1\}$ ,  $j \in \{1, \dots, M\}$  tetszőleges indexpár.

*Bizonyítás.* Tegyük fel, hogy az állítás nem igaz, vagyis  $\exists f : P_j^* \rightarrow P_i$  injektív függvény, amelyre  $l(f(y)) \geq l(y) \forall y \in P_j^*$  esetén. Először legyen  $i \in \{1, \dots, M\}$ . Legyen  $T' = T \setminus P_i$ . Ekkor a 7. Lemma teljesülése folytán az  $F$  algoritmus a  $T'$ -beli tárgyakat pontosan ugyanúgy helyezi el  $M-1$  ládába, mint az előbb, (és így jut egy tárgy az  $M$ -edik ládába is, amelyik az előbb az  $(M+1)$ -edik ládába került). Másrészt tekintsük a  $P^*$  pakolást, és konstruáljunk ebből egy  $P'$  pakolást úgy, hogy minden  $y \in P_j^*$  tárgyat cseréljünk ki az  $f(y)$  képével. Ekkor a  $P'_j$ -beli tárgyak mindegyike az  $P_i$  halmazhoz tartozik,  $l \neq j$  esetén a  $P'_l$ -beli tárgyak mérete csak csökkenhetett, ezért a  $P_i$ -beli tárgyakat elhagyva  $C^*[T', M-1] \leq q$ , másrészt  $F[T', p, M-1] = 0$ , ez pedig ellentmond annak, hogy a  $(T, M)$  pár minimális  $(p/q)$  ellenpélda. Most legyen  $i = M+1$ ,  $P_{M+1} = \{x\}$ . Ha most ez az  $M+1$ -edik láda dominálja a  $j$ -edik optimális ládát, akkor ebben az optimális ládában is csak egy tárgy van, legyen ez  $y$ , és ekkor  $l(x) \geq l(y)$ . Hagyjuk el a tárgyak közül  $x$ -et, vagyis legyen  $T' = T \setminus \{x\}$ . Ekkor az  $F$  algoritmus a  $T'$ -beli tárgyakat nem képes  $M-1$  ládába elhelyezni, mint ahogy az előbb sem tudta, és így jut az  $M$ -edik ládába is, pontosan azok, amelyek az előbb ebbe a ládába kerültek. Másrészt tekintsük a  $P^*$  pakolást, cseréljünk ki az  $x$  és  $y$  tárgyakat. Az  $x$  tárgy helyére nála kisebb  $y$  tárgy került,  $x$  pedig egyedül van a  $j$ -edik optimális ládában. Ezt a tárgyat elhagyva a többi befér  $M-1$  számú optimális ládába, vagyis újra azt kaptuk hogy

$C^*[T', M-1] \leq q$  és  $F[T', p, M-1] = 0$ , ami ellentmond annak, hogy a  $(T, M)$  pár minimális  $(p/q)$  ellenpélda.  $\square$

15. LEMMA.  $|P_j^*| \geq 2$ , minden  $1 \leq j \leq M$  esetén.

*Bizonyítás.* Ha  $P_j^* = \{y\}$  lenne valamilyen  $y \in T$  esetén, akkor ez az  $y$  tárgy benne van valamelyik  $P_i$  ládában  $i \in \{1, \dots, M+1\}$ , ekkor viszont  $P_i$  dominálja  $P_j^*$ -ot, ami ellentmond az előző Lemmának.

16. LEMMA. Legyen  $(T, M)$  tetszőleges  $(p/q)$  ellenpélda, legyen  $\alpha = l(T)$ , ahol  $T$  az első  $M$  ládából kimaradó tárgy. Ekkor  $\alpha > \frac{M}{M-1}(p-q)$ .

*Bizonyítás.* A  $T$  tárgy már nem fér be a  $P_i$  ládába, ezért  $l(P_i) + \alpha > p$  minden  $1 \leq i \leq M$  esetén, ezt  $i$ -re összegezve kapjuk:

$$(7) \quad \sum_{i=1}^M l(P_i) + M\alpha > Mp.$$

Másrészt

$$\sum_{i=1}^M l(P_i) + \alpha \leq Mq,$$

hiszen az optimális pakolás során minden tárgy befért az  $M$  számú  $q$  kapacitású ládába. Ezeket rendezve:

$$(M-1)\alpha > M(p-q),$$

ebből pedig a kívánt állítás adódik.  $\square$

17. TÉTEL.  $r_3(F_2) = \frac{15}{13}$ .

*Bizonyítás.* Álljon  $T$  a következő tárgyakból:  $T = \{7, 5, 5, 4, 3, 3, 3, 3, 3\}$ , legyen  $P_1^* = \{7, 3, 3\}$ ,  $P_2^* = \{5, 5, 3\}$ ,  $P_3^* = \{4, 3, 3, 3\}$ . Az  $F_2$  algoritmus általi pakolás  $15 - \varepsilon$  méretű ládába a következő lesz:  $P_1 = \{7, 5\}$ ,  $P_2 = \{5, 4, 3\}$ ,  $P_3 = \{3, 3, 3, 3\}$ , és egy 3 méretű tárgy kimarad a ládából. Ez azt jelenti, hogy  $r_3(F_2) \geq \frac{15}{13}$ .

Most belátjuk a másik irányú egyenlőtlenséget. Azt bizonyítjuk be, hogy  $(T, 3)$   $(15/13)$  ellenpélda nem létezik. Tegyük fel ezzel ellentétben, hogy  $(T, 3)$  minimális  $(15/13)$  ellenpélda az  $F_2$  algoritmus esetén. Az egyszerűbb jelölés kedvéért tételezzük fel, hogy az utolsó, a ládából kimaradó tárgyat a harmadik ládába tesszük, amelynek egy része így „kilóg” a ládából. Könnyen látható, hogy  $\min\{l(P_1), l(P_2)\} < 12$ , ugyanis  $l(T) \leq 39$ ,  $l(P_3) > 15$ , ezért az első két ládába összesen 24-nél kevesebb összméretű tárgy kerül. Legyen  $|P_1| = j$ ,  $|P_2| = l$ ,  $|P_3| = k$ . Jelöljük az első ládába kerülő tárgyakat  $A_i$ -vel ( $i = 1, \dots, j$ ), a másodikba kerülőket  $B_i$ -vel ( $i = 1, \dots, l$ ), a harmadikba kerülőket  $C_i$ -vel ( $i = 1, \dots, k$ ). (Az egyszerűség kedvéért a tárgyak hosszát is jelöljük  $l(A_i)$  helyett  $A_i$ -vel,  $l(B_i)$  helyett  $B_i$ -vel,  $l(C_i)$  helyett  $C_i$ -vel.) Mivel a harmadik ládába kerülő tárgyak az első kettő láda egyikébe sem férnek már be,  $C_i > 3$ . (Ugyanez a 16. Lemmából is következik.) Feltehető,

hogy az egyes ládákban lévő tárgyak méretei az indexsorrendjük szerint csökkennek, vagyis  $A_1 \geq A_2 \geq \dots \geq A_j$ , stb. Fel fogjuk használni a tétel bizonyításához a következő állításokat:

18. SEGÉDTÉTEL.  $|P_3| \leq 4$ .

*Bizonyítás.* Tegyük fel az állítás ellenkezőjét, vagyis azt, hogy  $|P_3| \geq 5$ . Legyen  $\sum_{i=1}^k C_i = 15 + \varepsilon$ , ahol  $\varepsilon > 0$ , hiszen az utolsó tárgy nem fér be  $P_3$ -ba. Mivel a tárgyak összmérete legfeljebb 39, az első két ládában együtt legfeljebb  $24 - \varepsilon$  összméretű tárgy van, ezért legalább egyikben az összméret legfeljebb  $12 - \frac{\varepsilon}{2}$ , így a  $P_3$ -beli tárgyak bármelyikének a mérete nagyobb, mint  $3 + \frac{\varepsilon}{2}$ , hiszen az előző ládákba nem férnek be. Ekkor  $\sum_{i=1}^k C_i > 5 \cdot \left(3 + \frac{\varepsilon}{2}\right) = 15 + \frac{5}{2}\varepsilon$ , ami ellentmond annak, hogy  $\varepsilon > 0$ .

□

19. SEGÉDTÉTEL. *Legfeljebb egy olyan optimális láda van, amelyben csak két tárgy van. Így a tárgyak száma legalább 8.*

*Bizonyítás.* Korábban láttuk, hogy minden optimális ládában legalább két tárgy van. Tegyük fel, hogy pontosan kettő tárgy van például a harmadik optimális ládában:  $X_3$  és  $Y_3$ , ahol  $X_3 \geq Y_3$ . Az  $F_2$  algoritmus a legelső ládába nem egy tárgyat helyez, mert akkor e tárgy mellé nem fér be másik, és akkor az optimális pakolásnál is egyedül lett volna. Ezért  $F_2$  a legelső ládába pontosan két tárgyat tesz be először, (aztán esetleg még néhányat). Ez az első két tárgy legyen  $X$  és  $Y$ . Ha ez a két tárgy az optimális pakolásnál egy ládában van, (ez legyen a  $j$ -edik optimális láda,  $P_j^*$ , ahol a  $j$  index lehet 1, 2 vagy 3 is), akkor a két legnagyobb  $P_j^*$ -ban levő tárgyról van szó. Ha  $P_j^*$ -ban nincs másik tárgy, akkor  $P_1$  dominálja  $P_j^*$ -ot, ami ellentmondás. Ha van másik tárgy is  $P_j^*$ -ban, akkor tegyük át a  $P_j^*$ -ban lévő többi tárgyat a harmadik ládába  $X_3$  és  $Y_3$  mellé. Mivel  $X + Y \geq X_3 + Y_3$ , elférnek. Ekkor megint azt kapjuk, hogy  $P_1$  dominálja  $P_j^*$ -ot, ellentmondás. Ezek szerint  $X$  és  $Y$  az optimális pakolásnál két különböző ládában van. Egyikük sem lehet a harmadik ládában, mert akkor a másik legalább akkora, mint a másik tárgy a harmadik ládában, és akkor  $P_1$  dominálja  $P_3^*$ -ot. Ezért például  $X$  az első,  $Y$  pedig a második optimális ládában van. Ha csak egy tárgy lenne valamelyik mellett, akkor megint azt kapnánk, hogy ezt a ládát  $P_1$  dominálja, ezért  $X$  mellett még legalább 2 tárgy van, hasonlóan  $Y$  mellett is. □

Az előző segédtétel általánosabban is igaz, a bizonyítása lényegében az előzőével megegyező:

20. SEGÉDTÉTEL. *Tetszőleges  $m \geq 3$  esetén legalább két olyan láda van, ahol a tárgyak száma legalább három.* □

21. SEGÉDTÉTEL.  $B_l \geq C_k$ , vagyis a második ládabeli legkisebb tárgy legalább akkora, mint a harmadik ládabeli legkisebb tárgy.

*Bizonyítás.* Tegyük fel, hogy  $B_l < C_k$ . Ekkor az  $F_2$  algoritmus a  $B_l$  tárgy helyett szívesebben tenné be a ládába a  $C_k$  tárgyat, azért nem tehette őt, mert nem

fér be a ládába, ezért  $C_k$  nagyobb, mint  $B_l$  és a második ládában megmaradó üres hely együttesen. Az első két ládában együtt legalább 6 méretű üres hely marad, az első ládabeli üres helynél  $B_l$  nagyobb, ezért  $B_l$  és a második ládában megmaradó üres hely együtt nagyobb, mint 6, így  $C_k > 6$ . A  $k$  szám nem lehet 1, 3-nál több sem lehet, mert már az első három tárgy sem fér be a ládába, továbbá nem lehet 3, mert akkor  $C_1 + C_2 > 12$ , és akkor inkább ez a két tárgy került volna abba a ládába az első kettő közül, ahol a tárgyak összmérete 12-nél kevesebb. Ezért  $k = 2$ . Amikor legfőljebb két tárgyat letettünk  $P_2$ -be, azok összmérete legalább  $C_1$ , emellé  $C_2$  nem fér be. Ezért  $|P_2| \leq 2$ , mert ha még lenne több tárgy (az első legfőljebb kettőn kívül)  $P_2$ -ben, azokat elhagyva is ellenpéldát kapnánk.  $|P_2|$  nem lehet 1, mert akkor az éppen a  $B_l$  tárgy, amelynek a helyére inkább  $C_k$  került volna. Ezért  $|P_2| = 2$ ,  $P_2 = \{B_1, B_2\}$ . Hagyjuk el  $B_2$ -t. Ekkor  $P_2$ -be vagy  $B_1$  kerül, és amellé semelyik  $P_3$ -beli tárgy nem fér és ellenpéldát kaptunk, vagy  $C_1$ , és akkor  $B_1$  és  $C_2$  megmarad a harmadik ládába, de oda nem férnek be, és ismét ellenpéldát kaptunk, vagyis az ellenpéldánk nem volt minimális. Ellentmondást kaptunk, ezért  $B_l \geq C_k$ .  $\square$

Ha  $|P_3| \leq 3$ , (vagyis  $k = 2$  vagy  $k = 3$ ), akkor legfőljebb két tárgyat elhelyezve  $P_1$ -be vagy  $P_2$ -be ezek összmérete legalább  $\sum_{i=1}^{k-1} C_i$ , ezek mellé a  $C_k$  tárgy nem fér be a ládába, de akkor az előző Segédttétel miatt ezek után semelyik  $P_2$ -beli és  $P_3$ -beli tárgy sem fér be az első-, és semelyik  $P_3$ -beli tárgy nem fér be a második ládába. Ezért a minimalitás miatt az első két ládában nincs is több tárgy, vagyis a tárgyak száma legfeljebb  $2 + 2 + 3 = 7$ , ami ellentmondás. Ezért, összevetve még a 18. Segédttétellel biztosan tudjuk, hogy  $|P_3| = 4$ .

Ha  $B_l < C_1$ , akkor az előző Lemma bizonyításával megegyező módon kapjuk, hogy  $C_1 > 6$ , és akkor a  $C_i$ -k száma legfeljebb 3, ami ellentmondás. Ezért  $B_l \geq C_1$ , és akkor az utolsó két ládában levő tárgyak méreteinek sorrendje a következő:  $B_1 \geq B_2 \geq \dots \geq B_l \geq C_1 \geq C_2 \geq C_3 \geq C_4$ .

## 22. SEGÉDTÉTEL. $|P_2| > 2$ .

*Bizonyítás.* Tegyük fel, hogy  $|P_2| \leq 2$ . Ez azt jelenti, hogy legfőljebb két tárgyat letéve a második ládába azok mellé  $C_4$  már nem fér be. Ekkor viszont ha legfőljebb kettő tárgyat először leteszünk az első ládába, azok összmérete legalább  $\sum_{i=1}^l B_i$ , ezért ezek mellé sem fér be  $C_4$ . Ekkor a minimalitás miatt az első ládában nincs is több tárgy, a tárgyak száma legfőljebb  $2 + 2 + 4 = 8$ . Mivel a tárgyak száma ennél kevesebb nem lehet, az első két ládában ekkor pontosan két-két tárgy van. Most tekintsük az optimális ládákat. A 19. Segédttételben leírt módon az előzőekből következik, hogy az első ládába letett két tárgy,  $A_1$  és  $A_2$  két különböző optimális ládában van, például az első két optimális ládában, és van még mellettük két-két tárgy, legyenek  $Z_1$  és  $Z_2$ , illetve  $Z_3$  és  $Z_4$ , és a harmadik optimális ládában van két tárgy:  $X_3$  és  $Y_3$ . (Ekkor persze  $T = \{A_1, A_2, Z_1, Z_2, Z_3, Z_4, X_3,$

$Y_3\} = \{A_1, A_2, B_1, B_2, C_1, C_2, C_3, C_4\}$ ,  $P_1^* = \{A_1, Z_1, Z_2\}$ ,  $P_2^* = \{A_2, Z_3, Z_4\}$ ,  $P_3^* = \{X_3, Y_3\}$ , valamint  $P_1 = \{A_1, A_2\}$ ,  $P_2 = \{B_1, B_2\}$ ,  $P_3 = \{C_1, C_2, C_3, C_4\}$ .)

Könnyen látható, hogy  $A_1 + A_2 > 11$ . Ha ugyanis  $A_1 + A_2 \leq 11$ , akkor a többi tárgy mindegyike nagyobb mint 4, mert az első ládába már nem férnek be. Ezért  $Z_1$  és  $Z_2$  bármelyike nagyobb, mint 4, ekkor  $A_1$  mérete legfőljebb  $13 - 4 - 4 = 5$ . Ekkor viszont  $A_1 + A_2 \leq 10$ , ebből az előzőhöz hasonló módon kapjuk hogy a többiek mérete legalább 5, ekkor  $A_1 \leq 3$ , ami ellentmondás, mert akkor  $A_1 + A_2 < 6 < 7.5 < C_1 + C_2$ . Ezért  $A_1 + A_2 > 11$ . Ekkor viszont  $Z_1 + Z_2 + Z_3 + Z_4 \leq 15$ , mert ez a négy tárgy  $A_1$  és  $A_2$ -vel együtt van az első két 13 méretű optimális ládában. Most azt vizsgáljuk meg, hogy melyik két tárgy lehet a  $Z_1, Z_2, Z_3, Z_4, X_3, Y_3$  tárgyak közül  $B_1$  és  $B_2$ . Ha két  $Z$  betűs, akkor ezek mellé még beférne a másik két  $Z$  betűs. Ha  $X_3$  és  $Y_3$ , akkor a maradék 4 tárgy befér a harmadik ládába. Ha  $X_3$  és  $Y_3$  egyike, és valamelyik  $Z$  betűs tárgy, akkor ezek ketten legalább  $X_3 + Y_3$  méretűek, a többi együtt legfőljebb  $Z_1 + Z_2 + Z_3 + Z_4$  méretű, és ezért beférnek a harmadik ládába. Más eset nem lehet, ellentmondást kaptunk.  $\square$

Ezek szerint  $|P_2| \geq 3$ . Viszont ha  $|P_2| \geq 4$  lenne, akkor a harmadik ládába beférne a négy  $C$  betűs tárgy. Ezért  $|P_2| = 3$ .

Ha  $l(P_1) > 12\frac{3}{4}$ , akkor  $l(P_2) < 11\frac{1}{4}$ , ekkor viszont  $l(P_3) < 11\frac{1}{4} \cdot \frac{4}{3} = 15$ , mert a harmadik ládában kisebb tárgyak vannak, mint a másodikban, és ellentmondást kapunk. Ezért  $l(P_1) \leq 12\frac{3}{4}$ . Tegyük le az első ládába legfeljebb két tárgyat. Ezek összmérete legalább  $C_1 + C_2 \geq \frac{1}{2} \cdot (C_1 + C_2 + C_3 + C_4) > 7.5$ . Ha most már nem fér be az első ládába  $P_2$ -, vagy  $P_3$ -beli tárgy, akkor a minimalitás miatt nincs több tárgy a  $P_1$  ládában. Ha fér még be, akkor legalább  $C_4$  méretével nő az összméret a  $P_1$  ládában, ezért legalább 3-mal. A kapott legalább 10.5 összméret mellé már nem fér be  $P_2$ , vagy  $P_3$ -beli tárgy, mivel a  $P_1$ -beli összméret legfeljebb  $12\frac{3}{4}$ . Így a tárgyhalmaz minimalitása miatt nincs több tárgy a  $P_1$  ládában. Azt kaptuk, hogy  $|P_1| \leq 4$ .  $|P_1|$  nem lehet 1, mert akkor ez az egy tárgy az optimális pakolásnál is egyedül lenne. Ezek után megvizsgáljuk a három lehetséges esetet az első ládában levő tárgyak száma szerint.

**4.3.1.**  $|P_1| = 2$ . Az optimális pakolásnál  $A_1$  nem lehet egyedül. Ha csak egy tárgy van mellette a ládájában, akkor feltehető, hogy ez a másik tárgy  $A_2$ , ekkor viszont a többi tárgy nem fér el a másik két ládában. Ezért  $A_1$  mellett legalább két tárgy van, és ezek egyike sem  $A_2$ . Ezek szerint  $A_1 \leq 13 - 3 - 3 = 7$ , a 15 méretű ládában legalább 8 méretű hely marad mellette. Könnyen látszik, hogy  $B_1$  nem lehet 8-nál nagyobb, mert akkor  $B_1 + B_2 + B_3 > 8 + 3 + 3 = 14$ , akkor  $A_1 + A_2 < 10$ , másrészt  $B_1 + B_2 > 8 + 3 = 11$ . Ezért  $A_1$  mellett bármelyik másik tárgy elfér, ebből következik, hogy  $A_1$  és  $A_2$  a két legnagyobb méretű tárgy, vagyis  $A_1 \geq A_2 \geq B_1 \geq B_2 \geq B_3 \geq C_1 \geq C_2 \geq C_3 \geq C_4$ . Mivel a négy legkisebb nem fér el egy 15 méretű ládában, mindegyik optimális ládában legfeljebb három, azaz pontosan három tárgy van. Legyen  $P_1^* = \{A_1, X_1, X_2\}$ ,  $P_2^* = \{A_2, Y_1, Y_2\}$ ,  $P_3^* = \{Z_1, Z_2, Z_3\}$ . (A tárgyak méretei az egyes optimális ládában monoton csökkenőek.) Most is belátható, hogy  $A_1 + A_2 > 11$ , és így  $X_1 + X_2 + Y_1 + Y_2 \leq 15$ .

Most azt vizsgáljuk, hogy melyik két tárgy kerül először a második ládába, vagyis mely tárgyak  $B_1$  és  $B_2$ .

1. Ha  $X_1$  és  $Y_1$ , akkor ezek mellé még  $X_2$  és  $Y_2$  elfér, ezért legalább  $X_1 + X_2 + Y_1 + Y_2$  összméret kerül a második ládába, de akkor az összes többi tárgy belefér a harmadik ládába. ellentmondás.
2. Ha  $Z_1$  és  $Z_2$ , akkor ezek mellé még  $Z_3$  elfér, legalább  $Z_1 + Z_2 + Z_3$  összméret kerül a második ládába, a többi tárgy pedig belefér a harmadik ládába, ellentmondás.
3. Ha  $Y_1$  és  $Z_1$ : Ha  $Y_1$  és  $Z_1$  legfeljebb 2-vel több, mint  $Z_1 + Z_2$ : Ezek mellé  $Z_3$  elfér, stb. Ha  $Y_1$  és  $Z_1$  több, mint 2-vel több, mint  $Z_1 + Z_2$ , akkor ezek mellé  $Y_2$  elfér, mert  $Z_1 + Y_1 \leq A_2 + Y_1$ . Ha most  $Y_2$  legfeljebb 2-vel kisebb, mint  $Z_3$ , akkor a második ládába megint legalább  $Z_1 + Z_2 + Z_3$  összméret kerül. Ellenkező esetben  $Z_3$  legalább 2-vel nagyobb, mint  $Y_2$ , ezért  $Z_3$  legalább 2-vel nagyobb, mint a legkisebb tárgy, így  $Z_3 > 5$ . Ekkor viszont  $Z_1 + Z_2 + Z_3 > 15$ , ellentmondás.
4. Ha  $X_1$  és  $Z_1$ : Az előző esethez hasonlóan kapunk ellentmondást.

**4.3.2.**  $|P_1| = 3$ . Ha  $A_3 < C_4$ , akkor  $A_3$  helyett bármelyik  $P_2$ , vagy  $P_3$ -beli tárgy jobb lenne, ezért ezek egyike sem férhet be az első ládába  $A_1$  és  $A_2$  mellé, ekkor viszont  $A_3$  elhagyható lenne, az ellenpélda nem minimális, ellentmondás. Ezért  $A_3 \geq C_4$ . Ekkor viszont  $A_3$  nagyobb, mint az üres hely az első ládában, és nagyobb, mint az üres hely a második ládában, mert ide  $C_4$  nem fért be.

Tegyük fel, hogy  $A_3 < B_1$ . Ekkor  $B_1$  jobb lenne  $A_3$  helyett, ezért nem fér be helyette a ládába, ezért  $B_1$  nagyobb, mint  $A_3$  és a mellette megmaradó üres hely, ez együtt nagyobb, mint az első két ládában megmaradó üres helyek mérete együtt, ami nagyobb, mint 6. Így  $B_1 + B_2 + B_3 > 6 + 3 + 3 = 12$ . Ekkor  $l(P_1) < 12$ . Legyen  $l(P_1) = 12 - \alpha$ , ahol  $\alpha > 0$ . Előbb láttuk, hogy  $B_1$  nagyobb, mint  $A_3$  és a mellette megmaradó üres hely. Az üres hely mérete pontosan  $3 + \alpha$ . Ennél  $A_3$  nagyobb, ezért  $B_1$  nagyobb, mint  $6 + 2\alpha$ . Ekkor  $l(P_2) > 6 + 2\alpha + 3 + 3 = 12 + 2\alpha$ , ami ellentmondás, mert akkor  $l(P_1) + l(P_2) > 24 + \alpha$ . Beláttuk, hogy  $A_3 \geq B_1$ . Ez viszont azt jelenti, hogy  $A_1, A_2$  és  $A_3$  a három legnagyobb tárgy, pontosabban  $A_1 \geq A_2 \geq A_3 \geq B_1 \geq B_2 \geq B_3 \geq C_1 \geq C_2 \geq C_3 \geq C_4$ . A négy legkisebb nem fér el egy 15 méretű ládában, ezért bármelyik optimális ládában is legfeljebb csak három tárgy lehet, ennél azonban nagyobb a tárgyak száma, ellentmondás.

*Megjegyzés.* A bizonyítás befejezése így is történhetett volna: Optimális ládában nem lehet legfeljebb három tárgy, mert akkor azt  $P_1$  dominálja, másrészt akkor legalább 3 · 4 tárgy kellene hogy legyen, ellentmondás.

**4.3.3.**  $|P_1| = 4$ . Tegyük fel, hogy  $l(P_1) < 11\frac{1}{4}$ . Ekkor a második és harmadik ládabeli tárgyak mérete nagyobb, mint  $3\frac{3}{4}$ . Másrészt  $A_1 + A_2 \geq C_1 + C_2 > 7.5$ . Így  $A_1$  és  $A_2$  mellé az első ládába nem fér be második vagy harmadik ládabeli tárgy, mert akkor az összméret az első ládában több lenne, mint  $7\frac{1}{2} + 3\frac{3}{4} = 11\frac{1}{4}$ . Ekkor a minimalitás miatt az első ládában csak két tárgy van, ellentmondás. Beláttuk, hogy  $l(P_1) \geq 11\frac{1}{4}$ , korábban pedig láttuk, hogy  $l(P_1) \leq 12\frac{3}{4}$ . Könnyen látható,



hogy  $C_4 \leq 4$ , mert különben a második és harmadik ládában lévő 7 tárgy együtt több mint 28 összméretű, és akkor  $l(P_1) < 11$  lenne.

Tegyük fel, hogy  $A_3 < C_4$ . Ekkor  $A_3$  helyett  $C_4$  jobb lenne, ezért nem fér be az első ládába helyette, ezért  $C_4$  nagyobb mint  $A_3$  és a mellette levő üres hely, így  $C_4$  több mint  $2\frac{1}{4}$ -del nagyobb mint  $A_3$ . Ezért  $A_3 < 4 - 2\frac{1}{4} = 1\frac{3}{4}$ . Ekkor  $A_3 + A_4 < 1\frac{3}{4} + 1\frac{3}{4} = 3\frac{1}{2}$ , másrészt  $A_3 + A_4 \geq B_1$ , ebből  $B_1 < 3\frac{1}{2}$  adódik. Azonban  $B_1 \geq C_1 > \frac{15}{4} = 3\frac{3}{4}$ , ellentmondás. Beláttuk, hogy  $A_3 \geq C_4$ .

Ha  $A_3$  legalább akkora lenne mint  $B_1$ , akkor  $A_1, A_2, A_3$  valamint a másik ládában lévő hét tárgy közül a négy legkisebb (vagyis  $C_1, C_2, C_4$  és  $C_4$ ) nem fér el egy ládában, így optimális ládában sem, de a számuk több, mint 9, így ellentmondást kapnánk. Ezért  $A_3 < B_1$ . Ekkor  $B_1$  jobb lenne  $A_3$  helyett, ezért nem fér be helyette a ládába, emiatt  $B_1 > A_3 + 2\frac{1}{4}$ , hiszen az első ládában az üres hely legalább  $2\frac{1}{4}$ . Kapjuk, hogy  $B_1 > 3 + 2\frac{1}{4} = 5\frac{1}{4}$ , mert  $A_3 \geq C_4 > 3$ . Másrészt  $A_3 + A_4 \geq B_1$ , így  $A_3 + A_4 \geq 5\frac{1}{4}$ . Ekkor azonban  $A_1 + A_2 + A_3 + A_4 > 7\frac{1}{2} + 5\frac{1}{4} = 12\frac{3}{4}$ , ami ellentmondás.

Ezzel beláttuk, hogy  $(T, 3)$  minimális (15/13) ellenpélda az  $F_2$  algoritmus esetén nem létezik, a Tétel bizonyítását befejeztük.  $\square$

*Megjegyzés.* Mivel a bizonyítás eléggé körülményes volt, nem törekszünk ezen a helyen arra, hogy további pontos becsléseket adjunk, azonban az előbbiek alapján azt sejtjük, hogy az  $F_2$  algoritmus esetében a pontos hatékonyságbecslés megegyezik az FFD algoritmus pontos hatékonyságbecslésével, vagyis  $4 \leq m \leq 7$  esetén  $r_m(F_2) = \frac{20}{17}$ , és nagyobb  $m$ -ekre is teljesül  $r_m(F_2) \leq 1.2$ . Hasonló okokból nem foglalkozunk e helyen az  $F_\gamma$  ( $\gamma \geq 3$ ) algoritmusokra vonatkozó pontos becslésekkel sem. Az előzőekből úgy tűnhet, hogy rögzített  $m$  esetén  $\gamma$  növelésével az  $r_m(F_\gamma)$  hatékonyság egyre javul, és a legjobb az  $F_\infty$  algoritmus esetén lesz. Ez nem igaz. Könnyen belátható például az alábbi

23. TÉTEL. Tetszőleges  $m > 3$  esetén  $r_m(F_2) \leq \frac{3}{2}$ .  $\square$

*Bizonyítás.* Tekintsünk egy  $(T, M)$  minimális (3/2) ellenpéldát. Legyenek az  $i$ -edik ládában az  $A_j^i$  ( $j = 1, \dots, |P_i|$ ) tárgyak  $i = 1 \dots m$  esetén, csökkenő méret szerint felsorolva. Könnyen látható, hogy az utolsó ládabeli tárgyak mindegyikének a súlya több mint 1, így legfölből három van belőlük. Legyen ezen tárgyak száma  $k$ , (ekkor  $k = 2$  vagy  $k = 3$ ). Legyen az  $i$ -edik ládabeli legkisebb tárgy mérete  $l_i$ , ( $i = 1, \dots, m$ ), ahol megint feltesszük hogy az utolsó ládában a legkisebb tárgy „kilóg” a ládából. Ezután belátjuk teljes indukcióval azt, hogy

1.  $|P_i| \leq 2$  ( $i = 1, \dots, m - 1$ ),
2.  $l_i \geq l_m$ , ( $i = 1, \dots, m - 1$ ).

Tegyük fel, hogy az állítást már beláttuk  $\alpha = i + 1, \dots, m - 1$  esetén, most belátjuk  $i$ -re. Amikor az  $F_2$  algoritmus elhelyezett legfölből 2 tárgyat az  $i$ -edik ládában, akkor ott ezek összsúlya legalább  $\sum_{j=1}^{k-1} A_j^m$ , ezért ezek mellé  $A_k^m$ , vagyis az utolsó ládából kilógó tárgy nem fér be, de akkor az indukciós feltétel miatt semelyik későbbi ládában levő tárgy sem. Ezért az ellenpélda minimalitása miatt az  $i$ -edik ládában nincs is több tárgy, vagyis 1-et beláttuk. Tegyük fel, hogy  $l_i < l_m$ . Ekkor az  $i$ -edik ládában pontosan kettő tárgy van, legyenek  $A$  és  $B$ . Hagyjuk el a

kisebbliket, vagyis  $B$ -t. Ekkor a maradék tárgyakra végrehajtva az  $F_2$  algoritmust a következő esetek lehetségesek:

a, Az  $i$ -edik ládába egyedül  $A$  kerül, és ekkor az összes további ládába ugyanazok a tárgyak, mint az előbb.

b, Az  $A$  tárgy kimarad az összes ládából. Ebben az eddig felsorolt két esetben kisebb elemszámú ellenpéldát kaptunk, vagyis az ellenpéldánk nem volt minimális és ellentmondást kaptunk. Az  $A$  tárgy mellé másik tárgy nem kerülhet, mert nem férnek be mellé.

c, Az előbbi eseten kívül ezért csak az lehetséges, hogy az  $A$  tárgy valamelyik későbbi ládába kerül. Ekkor az  $i$ -edik ládába néhány  $A$ -tól különböző tárgy kerül. Ezek pontosan azok a tárgyak lesznek, amelyek az előbb az  $(i+1)$ -edik ládában voltak. Ehhez hasonlóan addig a későbbi ládáig, amelyekben az  $A$  tárgy egymaga van, az  $A_j$  tárgyak eggyel előbbi ládába kerülnek, a későbbi ládában pedig pakolás ugyanaz, mint az előbb. Ebben az esetben is kisebb elemszámú ellenpéldát kaptunk, ellentmondáshoz jutottunk.  $\square$

Másrészt az alábbi példa mutatja, hogy  $r_7(F_\infty) > \frac{3}{2}$ .

*Példa.* Legyen  $m = 7$ , álljon  $\mathcal{T}$  hét darab 1.5 és hét darab 0.4 hosszúságú téglalaplóból. Ezek beférnek hét 1.9 méretű ládába, nyilván a következőképpen:  $P_i^* = \{1.5; 0.4\}$  ( $i = 1, \dots, 7$ ). Az  $F_2$  algoritmus általi pakolás  $3 - \varepsilon$  méretű ládába (ahol  $\varepsilon$  kicsi pozitív szám) a következő lesz: Az első ládába kerül mind a hét 0.4 méretű téglalap. A megmaradó hét darab 1.5 méretű téglalap pedig már nem fér el a maradék hat ládában, mert egy ládába csak egy fér el közülük. Ez azt jelenti, hogy  $r_7(F_\infty) \geq \frac{3}{1.9} > \frac{3}{2}$ .

## 5. Numerikus eredmények

Az  $F_\gamma$  és  $F_\infty$  algoritmusok numerikus vizsgálata során látni fogjuk, hogy sok esetben az előzőleg említett eredeti Multifit  $[K]$  algoritmusnál lényegesen jobban működő algoritmusokat kapunk. Számítógépes futtatásokat végeztünk, ahol Multifit-típusú algoritmusokat hasonlítottunk össze. A belső algoritmus az FFD, az  $F_2$ , az  $F_3$  illetve az  $F_\infty$  algoritmus volt. 100-szor futtattuk az algoritmusokat különböző feladatosztályokban, és azt számoltuk, hogy melyik algoritmus hányszor adott a négy között legjobb eredményt, illetve a  $\frac{C}{C_1}$  arányok átlagát, ahol  $C_1$  a (2)-beli alsó becslés,  $C$  pedig a heurisztikus megoldás értéke. A téglalapok hosszúságait egyenletes eloszlás szerint választottuk kerekítéssel a „par” oszlopában levő intervallumokból. Azt tapasztaltuk, hogy az új algoritmusok közül legalább az egyik legalább olyan „jó” volt száz futás eredményét tekintve mint az eredeti Multifit, számos esetben pedig lényegesen jobb legalább egyik új algoritmus.

Az első táblázatbeli esetekben  $\gamma$  növekedésével az algoritmus hatékonysága javul. Az első két paraméterosztály esetén  $F_\infty$  mindig optimális megoldást adott.

par	FFD	$F_2$	$F_3$	$F_\infty$	
1,16	89	92	98	100	min
1,16	1.003	1.002	1.001	1	$\frac{C}{C_1}$
5,20	24	40	63	100	min
5,20	1.017	1.010	1.006	1	$\frac{C}{C_1}$
20,35	8	10	23	100	$\frac{C}{C_1}$
20,35	1.017	1.010	1.013	1.0003	min
40,55	0	1	8	100	min
40,55	1.010	1.006	1.008	1.0002	$\frac{C}{C_1}$

1.  $m = 3, n = 15$  esete

par	FFD	$F_2$	$F_3$	$F_\infty$	
1,16	83	96	92	100	min
1,16	1.006	1.002	1.003	1	$\frac{C}{C_1}$
5,20	42	79	58	100	min
5,20	1.014	1.005	1.009	1.0002	$\frac{C}{C_1}$
20,35	0	29	0	99	min
20,35	1.056	1.024	1.056	1.011	$\frac{C}{C_1}$
40,55	0	39	0	98	min
40,55	1.057	1.034	1.057	1.028	$\frac{C}{C_1}$

2.  $m = 5, n = 19$  esete

Ahogy növekszik a tárgyak mérete, úgy lesz egyre jobb a többihez képest az  $F_\infty$  algoritmus.

A második esetben  $F_2$  és  $F_\infty$  a két legjobb algoritmus az előzőek közül. Érdekes módon az FFD és  $F_3$  algoritmusok esetén a  $[20,35]$  illetve  $[40,55]$  intervallumokból vett hosszúságok esetén a  $\frac{C}{C_1}$  arányok (nemcsak az első három tizedesjegyre) megegyeznek, ami arra utal, hogy az algoritmusok teljesen egyformán működnek.

Utoljára a következő paraméterekkel dolgoztunk: Véletlenszerűen választottuk meg a ládák  $m$  számát egyenletes eloszlás szerint, kerekítéssel a  $[3, 15]$  intervallumból. Ezután a tárgyak  $n$  számát egyenletes eloszlás szerint választottuk a  $[2m + 1, 5m]$  intervallumból. A tárgyak hosszát egyenletes eloszlás szerint választottuk a  $[20, 35]$  intervallumból, itt ugyanis az előbb már szignifikáns eltéréseket tapasztaltunk az egyes algoritmusok között. A futások eredményét az alábbi táblázat tartalmazza:

par	FFD	$F_2$	$F_3$	$F_\infty$	
20,35	5	19	20	91	min
20,35	1.052	1.037	1.036	1.015	$\frac{C}{C_1}$

## 3. eset

Látható, hogy az  $F_\infty$  algoritmus általában lényegesen jobban működik a többinél. A szerző ezúton köszöni meg Vizvári Béla értékes segítségét és tanácsait a feladat megfogalmazásában, és a cikk szerkezetének a kialakításában.

## Irodalom

- [1] E. E. Coffman, Jr., M. R. Garey and D. S. Johnson, An application of bin-packing to multiprocessor scheduling, *SIAM J. Comput.*, **7** (1978), 1–17.
- [2] Dósa Gy., Általánosított Multifit típusú módszerek egyforma párhuzamos gépek ütemezésére, *Alkalmazott Matematikai Lapok*, **19** (1999), 155–168.
- [3] D. K. Friesen, Tighter bounds for the multifit processor scheduling algorithms, *SIAM J. Comput.*, **13** (1984), 170–181.
- [4] D. K. Friesen and M. A. Langston, Evaluation of a MULTIFIT-based scheduling algorithm, *J. Algorithms*, **7** (1986), 35–59.
- [5] R. L. Graham, Bounds for certain multiprocessor anomalies, *Bell System Tech. J.*, **45** (1966), 1563–1581.
- [6] B. Vizvári and R. Demir, „It is Difficult to Find a Difficult Problem for the Scheduling of Identical Parallel Machines”, Department of Industrial Engineering of Bilkent University, Research Report, IEOR-9212, 1992.
- [7] B. Vizvári and R. Demir, „A Column Generation Algorithm to Schedule Identical Parallel Machines”, Rutgers University, Research Report, RRR-99-93, September 1994.
- [8] Vizvári B., *Bevezetés a termelésirányítás matematikai elméletébe*, Egyetemi jegyzet (ELTE, 1992).

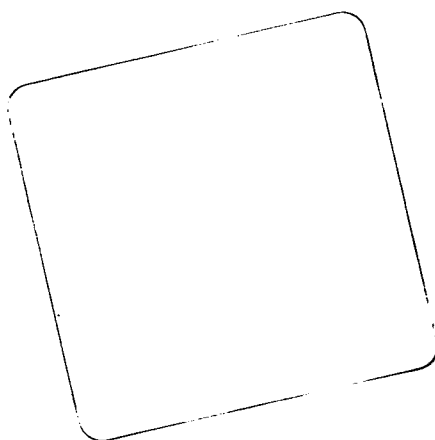
(Beérkezett: 1999. január 27.)

UNIVERSITY OF VESZPRÉM  
HUNGARY  
E-mail: DOSAGY@ALMOS.Vein.hu

## GENERALIZED MULTIFIT-TYPE METHODS II.

GYÖRGY DÓSA

We investigate a well known NP-complete problem of the scheduling theory: scheduling of identical parallel machines, that is: How to distribute  $n$  tasks among  $m < n$  machines to minimize the overall finishing time. In this article we generalize the classical algorithm Multifit: we change the bin-packing algorithm First Fit Decreasing with other bin-packing algorithms. We investigate the theoretical upper bounds of the new algorithms. Some numerical results show the efficiency of the new methods.



A kiadásért felelős a BJMT főtítkára  
Szedte az Egyenes Bt.

Nyomta az MSZH Nyomda és Kiadó Kft., Budapest  
Felelős vezető: Nagy László

Budapest, 2000  
Megjelent 10 (A/5) ív terjedelemben  
250 példányban  
HU ISSN 0133-3399

## ÚTMUTATÁS A SZERZŐKNEK

Az Alkalmazott Matematikai Lapok csak magyar nyelvű dolgozatokat közöl. A kéziratok gépelését olyan formában kérjük, hogy minden gépelt oldal 25, egyenként átlag 50 betűhelyes sort tartalmazzon. A közlésre szánt dolgozatokat három példányban kell beküldeni. Előnyben részesülnek a  $\text{\TeX}$ -ben elkészített dolgozatok. Ezeket két kinyomtatott példány kíséretében diszketten kérjük beadni.

A kéziratok szerkezeti felépítésének a következő követelményeket kell kielégíteni. A fejlécnek tartalmaznia kell a dolgozat címét, a szerző teljes nevét, valamint annak a városnak a nevét, ahol a szerző dolgozik. A fejléc után egy, képletet nem tartalmazó, legfeljebb 200 szóból álló kivonatot kell minden esetben megadni. A dolgozatot címmel ellátott szakaszokra kell bontani, és az egyes szakaszokat arab sorszámozással kell ellátni. Az esetleges bevezetésnek mindig az első szakaszt kell alkotnia. Az irodalomjegyzék után, a kézirat befejezésekképpen fel kell tüntetni a szerző teljes nevét és a munkahelye (illetve lakása) pontos címét. A dolgozatban előforduló képleteket szakaszonként újrakezdődően, a képlet előtt két zárójel közé írt kettős számozással kell azonosítani. Természetesen nem szükséges minden képletet számozással ellátni. Az esetleges definíciókat és tételeket (segédtevételeket és lemmákat) ugyancsak szakaszonként újrakezdődő, kettős számozással kell ellátni. Kérjük a szerzőket, hogy ezeket, valamint a tételek bizonyítását a szövegben kellő módon emeljék ki. Minden dolgozathoz csatolni kell egy angol, német francia vagy orosz nyelvű, külön oldalra gépelt összefoglalót. Amennyiben lehetséges, kérjük a nyomtatás számára különösen nehézkes matematikai jelölések használatának az elkerülését.

A dolgozatok ábráit és az esetleges lábjegyzeteket a dolgozat végén, különálló lapokon kérjük beküldeni. Mind az ábrákat, mind a lábjegyzeteket a dolgozat szakaszokra bontásától függetlenül, folytatólagos arab sorszámozással kell ellátni. Az ábrák elhelyezését a dolgozat megfelelő helyén, széljegyzetként feltüntetett, ábraazonosító sorszámokkal kell megadni. A lábjegyzetekre a dolgozaton belül az azonosító sorszám felső indexkénti használatával lehet hivatkozni.

Az irodalmi hivatkozások formája a következő. Minden hivatkozást fel kell sorolni a dolgozat végén található irodalomjegyzékben, a szerzők, illetve a társszerzők esetén az első szerző neve szerint alfabetikus sorrendben úgy, hogy a cirill betűs szerzők nevét a Mathematical Reviews átirási szabályai szerint latin betűsre kell átírni. A folyóiratban megjelent cikkekre [1], a könyvekre [5], a kötetben megjelent dolgozatokra [4], a disszertációkra [3] és a gépi program leírásokra [2] a következő minta szerint kell hivatkozni:

- [1] Farkas, J., Über die Theorie der einfachen Ungleichungen, *Journal für die reine und angewandte Mathematik* 124 (1902) 1–27.
- [2] Kéri, G., „DUALSIMP”, rutin a CDC 3300-ás gépekre (Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézete, CDC 3300 felhasználói ismertető 2. 1973. május) 19–20.
- [3] Prékopa, A., „Sztochasztikus rendszerek optimalizálási problémáiról”, doktori értekezés. Magyar Tudományos Akadémia, Budapest, 1970.
- [4] Prabhu, N. U., „Recent research on the ruin problem of collective risk theory”, in: *Inventory Control and Water Storage* Ed. A. Prékopa (János Bolyai Mathematical Society and North-Holland Publishing Company, Amsterdam–London, (1973) 221–228.
- [5] Zoutendijk, G. *Methods of Feasible Directions* (Elsevier Publishing Company, Amsterdam and New York, 1960).

A dolgozatok szövegében az irodalmi hivatkozás számait szögletes zárójelben kell megadni, mint például [5] vagy [4, 76–78]. A szerzők a dolgozatukról 50 darab ingyenes különlenyomatot kapnak. A dolgozatok után szerzői díjat az Alkalmazott Matematikai Lapok nem fizet.

## TARTALOMJEGYZÉK

<i>Mellár Tamás</i> , Az infláció és a külső-belső egyensúly együttmozgása .....	1
<i>Borgulya István</i> , Néhány polinomiális programozási feladat megoldása klaszterező algorit- mussal .....	19
<i>Kánnai Zoltán</i> , Dekompozábilis halmazok .....	35
<i>Bukszár József</i> , Multifákkal adott valószínűségi korlátok .....	45
<i>Kéri Gerzson és Rapcsák Tamás</i> , A talajvíz-szennyeződés modellezése és számítása Nagykáta térségében .....	61
<i>Pethőné Vendel Teréz</i> , Felületek előállítása trigonometrikusan súlyozott Spline-függvények felhasználásával .....	75
<i>Dósa György</i> , Általánosított multifit típusú módszerek II. ....	91

## INDEX

<i>Tamás Mellár</i> , Parallel movement of inflation, domestic and foreign equilibria .....	1
<i>István Borgulya</i> , Some polynomial programming problems solving with a clustering algo- rithm .....	19
<i>Kánnai Zoltán</i> , Decomposable sets .....	35
<i>József Bukszár</i> , Probability bounds given by multitrees .....	45
<i>Gerzson Kéri and Tamás Rapcsák</i> , Modelling of ground-water pollution in the area of Nagy- káta .....	61
<i>Teréz P. Vendel</i> , Flächengenerierung aus mit Trigonometrischen Binfunktionen Definierten Spline-Kurven .....	75
<i>György Dósa</i> , Generalized multifit-type methods II. ....	91



217.471 u ~~71~~

# Alkalmazott matematikai lapok

12

2000/2

A MAGYAR TUDOMÁNYOS AKADÉMIA  
MATEMATIKAI TUDOMÁNYOK  
OSZTÁLYÁNAK KÖZLEMÉNYEI

20.

KÖTET

# ALKALMAZOTT MATEMATIKAI LAPOK

## A MAGYAR TUDOMÁNYOS AKADÉMIA MATEMATIKAI TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

ALAPÍTOTTÁK

KALMÁR LÁSZLÓ, TANDORI KÁROLY, PRÉKOPA ANDRÁS, ARATÓ MÁTYÁS

FŐSZERKESZTŐ

BENCZÚR ANDRÁS

FŐSZERKESZTŐ-HELYETTESEK

DEMETROVICS JÁNOS, FARKAS MIKLÓS

FELELŐS SZERKESZTŐ

SZÁNTAI TAMÁS

A SZERKESZTŐBIZOTTSÁG TAGJAI

Arató Mátyás, Csirik János, Csiszár Imre, Galántai Aurél, Gécseg Ferenc, Gyires Béla, Györfy László, Harnos Zsolt, Hatvani László, Heppes András, Kátai Imre, Katona Gyula, Kis Ottó, Klafszy Emil, Kovács Margit, Lovász László, Maros István, Prékopa András, Recski András, Stoyan Gisbert, Szentkúti Zsolt (technikai szerkesztő), Tandori Károly, Tusnády Gábor, Varga László

XX. kötet 2. szám

Szerkesztőség és kiadóhivatal: 1027 Budapest, Fő u. 68.

Az Alkalmazott Matematikai Lapok változó terjedelmű füzetekben jelenik meg, és olyan eredeti tudományos cikkeket publikál, amelyek a gyakorlatban, vagy más tudományokban közvetlenül felhasználható új matematikai eredményt tartalmaznak, illetve már ismert, de színvonalas matematikai apparátus újszerű és jelentős alkalmazását mutatják be. A folyóirat közöl cikk formájában megírt, új tudományos eredménynek számító programokat, és olyan, külföldi folyóiratban már publikált dolgozatokat, amelyek magyar nyelven történő megjelentetése elősegítheti az elért eredmények minél előbbi, széles körű hazai felhasználását. A szerkesztőbizottság bizonyos időnként lehetővé kívánja tenni, hogy a legjobb cikkek nemzetközi folyóiratok különszámaként angol nyelven is megjelenhessenek.

A folyóirat feladata a Magyar Tudományos Akadémia III. (Matematikai) Osztályának munkájára vonatkozó közlemények, könyvismertetések stb. publikálása is.

A kéziratok a főszerkesztőhöz, vagy a szerkesztőbizottság bármely tagjához beküldhetők. A főszerkesztő címe:

Benczúr András, főszerkesztő  
1027 Budapest, Fő u. 68.

Közlésre el nem fogadott kéziratokat a szerkesztőség lehetőleg visszajuttat a szerzőhöz, de a beküldött kéziratok megőrzéséért vagy továbbításáért felelősséget nem vállal.

Az Alkalmazott Matematikai Lapok előfizetési ára kötetenként 850 forint. Megrendelések a szerkesztőség címén lehetségesek.

A Magyar Tudományos Akadémia III. (Matematikai) Osztálya a következő idegen nyelvű folyóiratokat adja ki:

1. Acta Mathematica Hungarica,
2. Acta Physica Hungarica.

## FUZZY DÖNTÉSI ELJÁRÁSOK ÖSSZEHAISONLÍTÁSA IRÁNYÍTÁSTECHNIKAI SZEMPONTBÓL

SCHUSTER GYÖRGY

Budapest

A fuzzy logika megjelenése óta számos döntési eljárást alkalmaztak különböző folyamatok irányítására. Ezen eljárások bizonyos mértékben eltértek egymástól, és mind-egyiknek megvannak az előnyös tulajdonságai, illetve hátrányai. Ebben a cikkben öt döntési eljárás összehasonlítása történik meg, átviteli karakterisztikáik, illetve érzékenyséjük alapján. A cikk második részében ezen eljárások felhasználásával egy gyakorlati példán is bemutatjuk az eljárások közti különbséget.

### 1. Bevezetés

A fuzzy halmazokat Lotfi A. Zadeh vezette be 1965-ben [6] mint egy módszert, amellyel a nem pontosan rendelkezésre álló adatokat értelmezni, ábrázolni és kezelni tudjuk.

*1.1. Definíció.* Legyen  $X$  egy nem üres halmaz. Az  $X$  halmaz egy  $A$  fuzzy részhalmazát a  $\mu_A$ -val jelölt tagsági függvényével írjuk le, ahol

$$\mu_A : X \rightarrow [0, 1]$$

és  $\mu_A(x)$ -et úgy interpretáljuk, mint az  $x$  elemnek az  $A$ -ba való tartozásának a mértékét, minden  $x \in X$  esetén. Ha valamely  $y \in X$  és  $\mu_A(y) = 0$ , akkor azt mondjuk, hogy  $y$  nulla mértékkel tartozik az  $A$  fuzzy halmazba. Ha  $\mu_A(y) = 1$ , akkor azt mondjuk, hogy  $y$  teljes mértékkel tartozik az  $A$  fuzzy halmazba. Nagyon sokszor  $\mu_A(x)$  helyett egyszerűen csak  $A(x)$ -et fogunk írni. Az  $X \neq \emptyset$  fuzzy halmazainak a családját  $\mathcal{F}(X)$ -el, míg az  $\mathbf{R}$  fuzzy halmazainak a halmazát  $\mathcal{F}$ -el jelöljük.

A fuzzy halmazok olyan tulajdonságok leírására alkalmasak, amelyeket nem lehet jellemezni a hagyományos eleme relációval, illetve nem kezelhetők a klasszikus logikai igen/nem segítségével.

*1.2. Definíció.* Az egyelemű halmazok karakterisztikus függvényét (fuzzy) szingletonnak nevezzük. Ha a fuzzy szingleton tartója az az  $x_0 \in \mathbf{R}$  pont akkor az  $\{x_0\}$  halmaz karakterisztikus függvényét  $\bar{x}_0$ -al fogjuk jelölni.

## 1.3. Definíció. A

$$T : [0, 1] \times [0, 1] \rightarrow [0, 1]$$

függvényt trianguláris nnormának (röviden t-normának) hívjuk, ha  $T$  szimmetrikus, asszociatív, monoton növekvő mindkét változójában és  $T(a, 1) = a$  minden  $a \in [0, 1]$  esetén.

A t-normákat a logikai és kapcsolat modellezésére használjuk. A cikkben két alapvető t-normát használunk, a minimumot és a szorzatot.

## 1.4. Definíció. Az

$$S : [0, 1] \times [0, 1] \rightarrow [0, 1]$$

függvényt t-konormának hívjuk, ha  $S$  szimmetrikus, asszociatív, monoton növekvő mindkét változójában és  $S(a, 0) = a$  minden  $a \in [0, 1]$  esetén.

A t-konormákat a logikai vagy kapcsolat modellezésére használjuk. A cikkben két alapvető t-konormát alkalmazunk, a maximumot és a normált szummát.

1.5. Definíció. Az  $A$  fuzzy halmaz normális, ha létezik olyan  $a \in X$ , hogy  $\mu_A(a) = 1$ . Ellenkező esetben az  $A$ -t szubnormálisnak hívjuk.

1.6. Definíció. Egy nyelvi változó a következő ötessel definiálható:

$$(x, T(x), U, G, M)$$

ahol  $x$  a változó neve,  $T(x)$  azoknak az elnevezéseknek a halmaza, amelyeket az  $x$  felvehet,  $U$  az  $x$  értélkeinek az értelmezési tarománya,  $G$  a formai szabály az  $x$  elnevezéseinek megadására, és  $M$  a tartalmi szabály az elnevezések tagsági függvényeinek megadására.

Például egy hőmérsékleti skála a következőképpen adható meg:

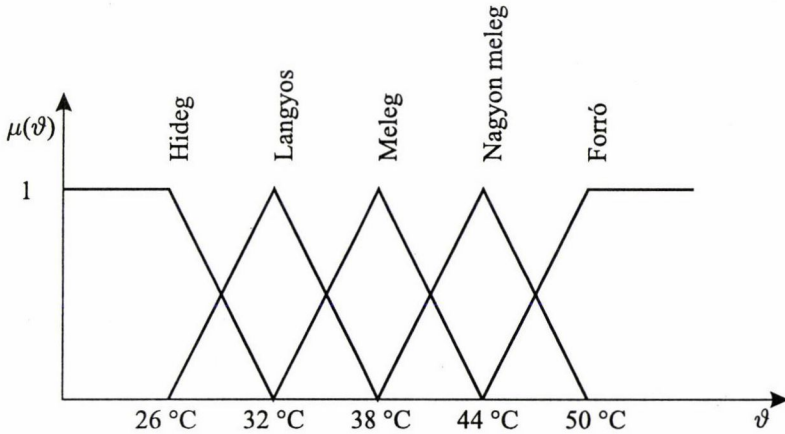
$$T(x) = \{\text{hideg, langyos, meleg, nagyon\_meleg, forró}\}$$

ahol  $x$  a hőmérsékletet jelenti, mint nyelvi változó. Nyilván, a felsorolt nyelvi értékek másképp vannak definiálva a ha a környezet változik, pl. a *forró* mást jelent a szaunában mint a tengerparton.

Egy tipikus fuzzy szabálybázis a következőképpen adható meg:

$$\begin{array}{llll} \mathcal{R}_1 : & \text{Ha „}x \text{ az } A_1\text{” és „}y \text{ az } B_1\text{” akkor} & \text{„}z \text{ az } C_1\text{”} & \text{vagy} \\ \mathcal{R}_2 : & \text{Ha „}x \text{ az } A_2\text{” és „}y \text{ az } B_2\text{” akkor} & \text{„}z \text{ az } C_2\text{”} & \text{vagy} \\ \vdots & & & \text{vagy} \\ \mathcal{R}_n : & \text{Ha „}x \text{ az } A_n\text{” és „}y \text{ az } B_n\text{” akkor} & \text{„}z \text{ az } C_n\text{”} & \end{array}$$

ahol, az  $A_i, B_i \in \mathcal{F}$  fuzzy halmazok jelölik az  $x$  és  $y$  nyelvi változó adott értékeit értelemszerűen, és  $C_i \in \mathcal{F}$  jelöli a  $z$  nyelvi változó adott értékét,  $i = 1, \dots, n$ . Az



1. ábra: Nyelvi szabályok

és kapcsolatot egy megfelelő t-norma, a **vagy** kapcsolatot egy célszerűen megválasztott t-konorma, míg a **ha...akkor** kapcsolatot egy megfelelően választott fuzzy implikáció reprezentálja. Az „ $x$  az  $A_1$ ” típusú fuzzy állítások teljesülési mértékét pedig pontosan az  $A_1$  tagsági függvényével definiáljuk.

A fuzzy döntési eljárások alatt olyan következtetési módszereket értünk, amelyekben az  $x_0 \in \mathbf{R}$  és  $y_0 \in \mathbf{R}$  (determinisztikus) inputokból és az

$$\mathcal{R} = \{\mathcal{R}_1, \dots, \mathcal{R}_n\}$$

szabálybázisból egy (determinisztikus)  $z_0 \in \mathbf{R}$  fuzzy outputot származtatunk. A fuzzy döntési eljárások működése öt lépésben foglalható össze:

1. A determinisztikus inputok fuzzifikálása.
2. A szabályok illetékességi szintjének a meghatározása.
3. Az egyes szabályok outputjának a meghatározása.
4. A (fuzzy) rendszer output meghatározása a szabályok outputjainak az aggregálása révén.
5. A kapott (fuzzy) halmaz egy, ugynevezett legjellemzőbb pontjának a kiválasztása (defuzzifikáció).

A Zadeh által 1973-ban bevezetett [9] kompozíciós következtetési szabály a leggyakrabban használt következtetési szabály a fuzzy szabálybázis alapu rendszerekben. A fuzzy döntéseket a következtetés kompozíciós szabálya segítségével hozzuk meg a következőképpen:

$\mathfrak{R}_1$ :	Ha „ $x$ az $A_1$ ” és „ $y$ az $B_1$ ” akkor	„ $z$ az $C_1$ ”	vagy
$\mathfrak{R}_2$ :	Ha „ $x$ az $A_2$ ” és „ $y$ az $B_2$ ” akkor	„ $z$ az $C_2$ ”	vagy
$\dots \&$			vagy
$\mathfrak{R}_n$ :	Ha „ $x$ az $A_n$ ” és „ $y$ az $B_n$ ” akkor	„ $z$ az $C_n$ ”	
input :	$x_0$ és $y_0$		
<hr/>			
output :		$z_0$	

ahol a  $z_0$  rendszer outputot a következő általános képlet segítségével számoljuk ki

$$z_0 = \text{defuzz}[C].$$

ahol

$$C = \text{Agg}(\text{fact} \circ \mathfrak{R}_1, \dots, \text{fact} \circ \mathfrak{R}_n) = \text{Agg}(\bar{x}_0 \times \bar{y}_0 \circ \mathfrak{R}_1, \dots, \bar{x}_0 \times \bar{y}_0 \circ \mathfrak{R}_n)$$

jelöli a rendszer fuzzy outputját, amit pontonként a

$$C(w) = \text{Agg}\{A_1(x_0) \times B_1(y_0) \rightarrow C_1(w), \dots, A_n(x_0) \times B_n(y_0) \rightarrow C_n(w)\}, \quad w \in \mathbf{R},$$

formulával számolunk ki. A fuzzy döntési eljárás öt lépése tehát így foglalható össze:

1. A determinisztikus inputok fuzzifikálása fuzzy szingletonok segítségével.
2. Az  $i$ -dik szabályok illetékességi szintjének a meghatározása a

$$A_i(x_0) \times B_i(y_0)$$

formulával, ahol a  $\times$  egy jól választott  $t$ -normát jelöl [2].

3. Az  $i$ -dik szabály outputjának a meghatározása a

$$C'_i(w) := A_i(x_0) \times B_i(y_0) \rightarrow C_i(w), \quad w \in \mathbf{R},$$

formulával, ahol  $\rightarrow$  egy jól választott fuzzy implikációt jelöl [2].

4. A rendszer (fuzzy) output meghatározása a

$$C = \text{Agg}\{C'_1, \dots, C'_n\}$$

formulával, ahol Agg egy jól választott aggregáló operátort jelöl [2].

5. A kapott (fuzzy) halmaz egy, ugynevezett legjellemzőbb pontjának a kiválasztása (defuzzifikáció) a

$$z_0 = \text{defuzz}[C].$$

formulával, ahol defuzz egy megfelelően választott defuzzifikáló operátor [2].

A döntések ábrázolásának egyik módja a FAM (Fuzzy Association Map) táblák alkalmazása, ahol a táblázat élein a bemeneti szabályokat a táblázat belsejében az implikált kimeneti szabályokat adjuk meg.

## 2. Gyakorlati megfontolások

A dolgozatban a következő döntési eljárásokat vizsgáljuk:

1. A minimum-maximum-súlypont (*min-max-gravity*) (MMG) módszer, úgynevezett Mamdani-féle döntés [2],
2. A szorzás-maximum-súlypont (*prod-max-gravity*) (PMG) módszer,
3. A szorzás-összeadás-súlypont (*prod-sum-gravity*) (PSG) módszer,
4. A szorzás-összeadás-szingletonokkal (*singleton-used-prod-sum-gravity*) (SPSG) módszer [3],
5. A Tsukamoto-féle döntési módszer [5].

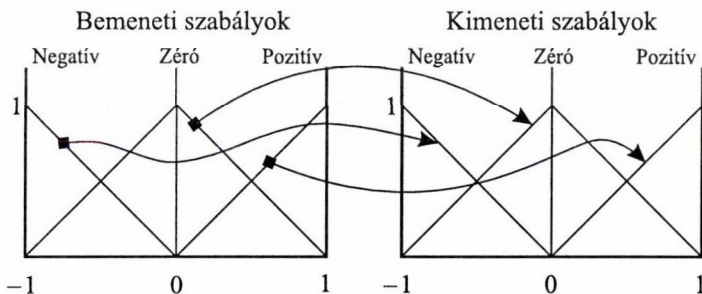
*2.1. Megjegyzés.* A gyakorlatban nagyon elterjedt úgynevezett *interpolációs típusú* fuzzy döntési eljárások [7, 8] tesztelésével és összehasonlításával (egy négyke-rék kormányzású robottargonca irányításának szimulációja tapasztalatai alapján) egy soron következő munkában foglalkozunk.

A fuzzy szabályrendszerekben szereplő szabályok felvétele nagyrészt önkényes, vannak ugyan szisztematikus módszerek a szabályok alakjának és helyének meghatározására, de ezek alkalmazása nem minden esetben lehetséges, mert a feladat nem teszi lehetővé a szabályok módszeres előzetes vizsgálatát.

Ezért a különböző döntési eljárások összehasonlítása nehézségekbe ütközik. Ennek a problémának a feloldásra a következő szabályokat vezettük be:

1. A szabályrendszerek minden döntési eljárás során azonos alakúak (lásd 2. ábra). Ez alól kivétel *singleton-used-prod-sum-gravity* módszer, ahol a kimeneti szabályrendszer szingletonokból áll.
2. A bemeneti fuzzy változó minden esetben szingleton.
3. A szabályrendszerek szimmetrikusak, így az átviteli karakterisztika minden vizsgált esetben páratlan függvény, ezáltal elegendő csak a karakterisztika egyik oldalát vizsgálni.
4. A bementi és kimeneti szabály párok azonos alakúak, így kezelésük könnyebb, nem kell zavaró konstansokkal foglalkozni. Bár így elveszítjük az általánosság megszorítását, azonban a gyakorlati tapasztalatok teljes mértékben alátámasztják a vizsgálat eredményeit.
5. Csak olyan fuzzy-crisp átalakítást vizsgálunk, amely súlypont számításra alapul [1].





2. ábra: Bemeneti és kimeneti szabályrendszerek

**2.2. Megjegyzés.** Amennyiben nem súlypont módszeren alapuló fuzzy-crisp átalakítást alkalmazunk az összehasonlító vizsgálat nagyon nehezzé válik, mivel az átviteli karakterisztika általában nem folytonos függvény, továbbá irányítástechnikai alkalmazásra csak igen korlátozott mértékben használható döntési eljárást kapnánk.

### 3. Döntési eljárások

Az döntési algoritmusok vizsgálatánál alkalmazott szabályok a következők (lásd 2. ábra)

$$\begin{aligned} \mathcal{R}_1 : & \text{Ha „}x \text{ az pozitív”} && \text{akkor „}y \text{ az pozitív”} && \text{vagy} \\ \mathcal{R}_2 : & \text{Ha „}x \text{ az zero”} && \text{akkor „}y \text{ az zero”} && \text{vagy} \\ \mathcal{R}_3 : & \text{Ha „}x \text{ az negatív”} && \text{akkor „}y \text{ az negatív”} \end{aligned}$$

ahol a szabálybázisban szereplő fuzzy halmazok tagsági függvényei a következők

$$\begin{aligned} \mu_{\text{negatív}}(u) &= \begin{cases} -u & \text{ha } -1 \leq u \leq 0 \\ 0 & \text{különben} \end{cases} \\ \mu_{\text{zero}}(u) &= \begin{cases} 1 - |u| & \text{ha } -1 \leq u \leq 1 \\ 0 & \text{különben} \end{cases} \\ \mu_{\text{pozitív}}(u) &= \begin{cases} u & \text{ha } 0 \leq u \leq 1 \\ 0 & \text{különben} \end{cases} \end{aligned}$$

Az egyszerűség kedvéért ugyanazt a fuzzy partíciót alkalmaztuk az input és az output tér esetén is.

Most megvizsgáljuk, hogy mi lesz a defuzzifikált output a különböző típusú döntési módszerek alkalmazása esetén. A döntési eljárások esetén csak az  $x \geq 0$  esetet vizsgáljuk, mivel az átviteli karakterisztikák mind páratlan függvények.

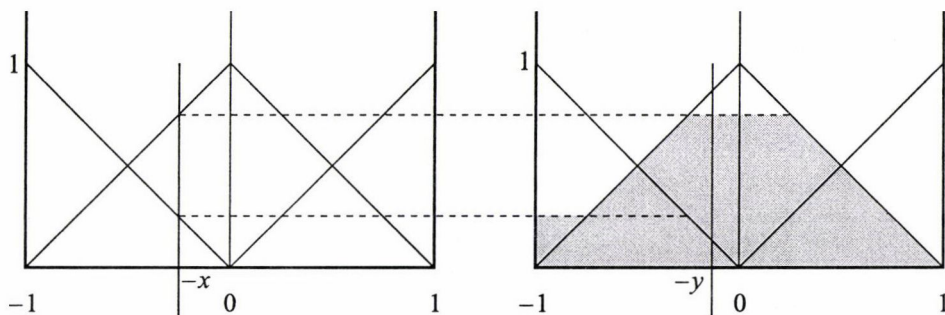


### 3.1. MMG döntés

A döntés működése a következő: a bemeneti jelet szingletonná alakítjuk ezzel a szingletonnal elmetsszük az érintett szabályokat, majd ezeket a metszeti értékekkel minimum kapcsolatba hozzuk a hozzárendelt kimeneti szabályokat (lásd 2. ábra). Az így kapott szabályokat maximum operátor segítségével egyesítjük. Az eredményként kapott fuzzy halmaz súlypontjának vízszintes tengelyre vett vetülete a döntés crisp kimenete.

Az átviteli karakterisztika számításához ezt a döntési eljárást két részre kell bontanunk az abszolútérték függvény miatt a)  $x \leq 0,5$  és b)  $x > 0,5$ .

Az a) esetben a döntést a 3. ábra mutatja.



3. ábra: Min-max-gravity döntés  $x \leq 0,5$  esetben

A karakterisztika a következő kifejezéssel számítható:

$$y_s(x) = \frac{\int_{-1}^{-x} y(1+y) dy + \int_{-x}^x y(1-x) dy + \int_x^{1-x} y(1-y) dy + \int_{1-x}^1 yx dy}{\int_{-1}^{-x} (1+y) dy + \int_{-x}^x (1-x) dy + \int_x^{1-x} (1-y) dy + \int_{1-x}^1 x dy}$$

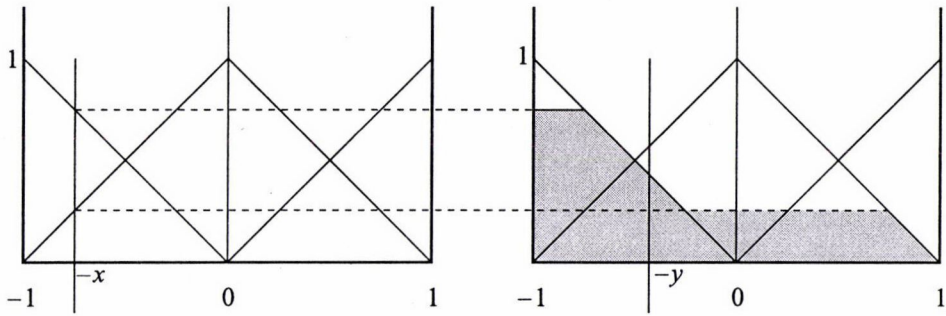
$$= \frac{1}{3} \times \frac{1-x^3+3x^2}{-x^2+2}.$$

A b) esetet a 4. ábra mutatja.

A karakterisztikát leíró kifejezés:

$$y_s(x) = \frac{\int_{-1}^{-x} y(1+y) dy + \int_{-x}^{1-x} y(1-x) dy + \int_{1-x}^x y(1-y) dy + \int_x^1 yx dy}{\int_{-1}^{-x} (1+y) dy + \int_{-x}^{1-x} (1-x) dy + \int_{1-x}^x (1-y) dy + \int_x^1 x dy}$$

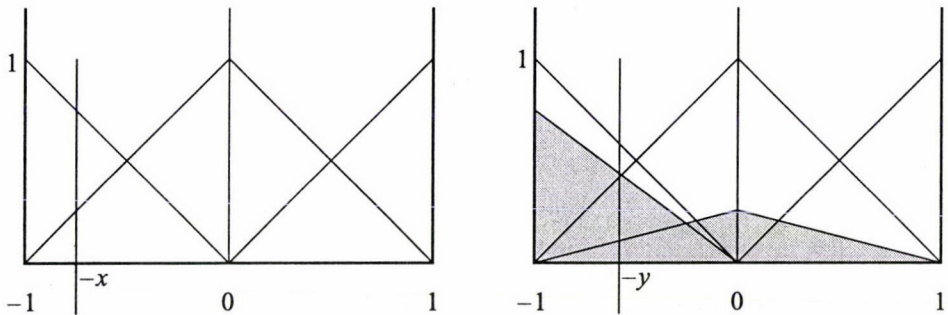
$$= \frac{1}{3} \times \frac{1-x^3+3x^2}{-x^2+2}.$$

4. ábra: Min-max-gravity döntés  $x > 0,5$  esetben

Érdekes módon az tapasztalható, hogy a két esetben a karakterisztikát leíró kifejezés azonos. Ennek oka a szabályrendszerek alakjában keresendő.

### 3.2. PMG döntés

Ez a döntési eljárás nagyon hasonló az előző módszerhez, de a kimeneti szabályokon alkalmazott minimum kapcsolat helyett a szabályokat szorozzuk a megfelelő metszeti értékekkel, majd ezeket a részeredményeket maximum kapcsolatba hozzuk. A kimenet itt is az eredő halmaz súlypontjának vízszintes tengelyre vett vetülete. A döntés működését a 5. ábra mutatja:



5. ábra: Prod-max-gravity döntés

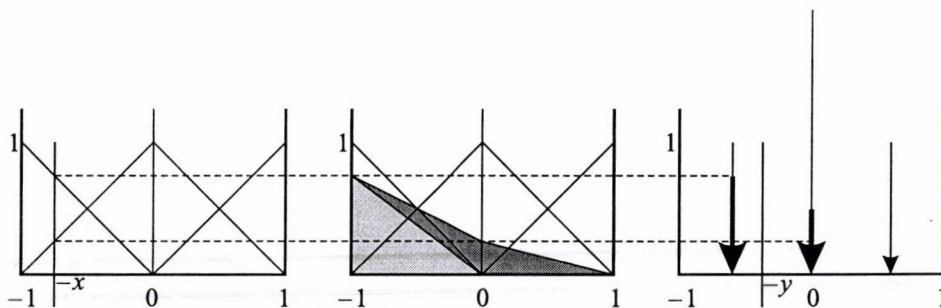
A karakterisztikát a következő kifejezés írja le:

$$\begin{aligned}
 y_s(x) &= \frac{\int_{-1}^0 y(1-x)(1+y) dy + \int_0^{1-x} y(1-x)(1-y) dy + \int_{1-x}^1 xy dy}{\int_{-1}^0 (1-x)(1+y) dy + \int_0^{1-x} (1-x)(1-y) dy + \int_{1-x}^1 xy dy} \\
 &= \frac{1}{3} \times \frac{x^4 - 2x^3 + 3x^2}{x^2 - 2x + 2}.
 \end{aligned}$$

### 3.3. PSG döntés

Ez a döntési eljárás az előzőtől abban különbözik, hogy a kimeneti részdöntések egyesítését nem maximum kapcsolat segítségével végzi, hanem a részdöntés szabályait összeadja. A kimeneti érték itt is súlypontszámítással történik.

A PSG döntés első pillanatra nagyon hasonló az előző döntésekhez, azonban az összeadás – ellentétben a maximum operátorral – nem takar el területrészeket a szabályokból. Továbbá a szorzás miatt egy részdöntés szabályának területe megegyezik az eredeti szabály területének és a megfelelő metszeti értéknek a szorzatával. Ezért a döntés számításánál nem kell a teljes kimeneti szabályrendszert végigszámolni, elegendő az tudni, hogy mekkora az adott szabály területe és a vízszintes tengelyen hol helyezkedik el a súlypontja. Így csak egy egyszerű néhány pontból álló nyomatéki egyenletet kell megoldani, amely radikálisan csökkenti a döntést megvalósító algoritmus futási idejét az előző eljárásokhoz képest.



6. ábra: Prod-sum-gravity döntés

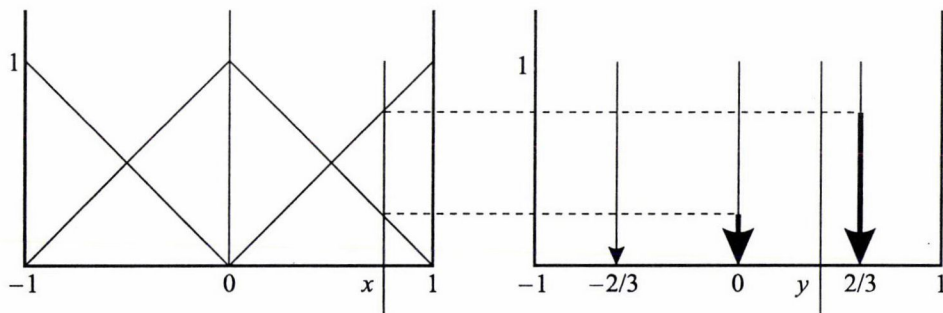
A karakterisztikát leíró kifejezés:

$$y_s(x) = \frac{\int_{-1}^0 y(1-x)(1+y) dy + \int_0^1 y(1-x)(1-y) dy + \int_0^1 yxy dy}{\int_{-1}^0 (1-x)(1+y) dy + \int_0^1 (1-x)(1-y) dy + \int_0^1 xy dy}$$

$$= \frac{2x}{6-3x}.$$

### 3.4. SPSG döntés

Ebben az esetben a kimeneti szabályokat szingletonok helyettesítjük, a szingletonok helye a könnyebb összehasonlítás érdekében az eredeti szabályok súlypontjaiban van. Ez az eljárás nagyon egyszerűvé teszi a crisp döntés kiszámítását. A 7. ábra ennek működését mutatja.



7. ábra: Singleton-used-prod-max-gravity döntés

A karakterisztikát leíró kifejezés:

$$y_s(x) = \frac{2x}{3}.$$

### 3.5. Tsukamoto-féle döntés

A Tsukamoto-féle döntés esetén az algoritmus jelentősen eltér az előzőekben látott döntésektől, ebben az esetben a bemeneti – kimeneti szabály párok egy függvény – inverz függvény kapcsolatot valósítanak meg. Tekintsük a 8. ábrát.

Vegyük példának az a) be–ki szabáлыпárt. Az adott bemeneti  $x$  érték egy metaszeti szintet jelöl ki a bemeneti szabályon. Ezt a szintet átvetítjük a kimeneti szabályra és meghatározzuk, hogy milyen kimeneti értéket határoz meg. A döntés működése miatt a következő szempontokat kell szem előtt tartani:

- a bemeneti szabályt leíró függvény célszerű ha monoton, mert így maga a szabály nem okoz stabilitási problémákat,
- a kimeneti szabályt leíró függvénynek szigorúan monotonnak kell lennie.

Az adott esetben a második feltétel úgy teljesíthető, hogy a kimeneti szabályok értelmezési tartományát a következő módon határozzuk meg:

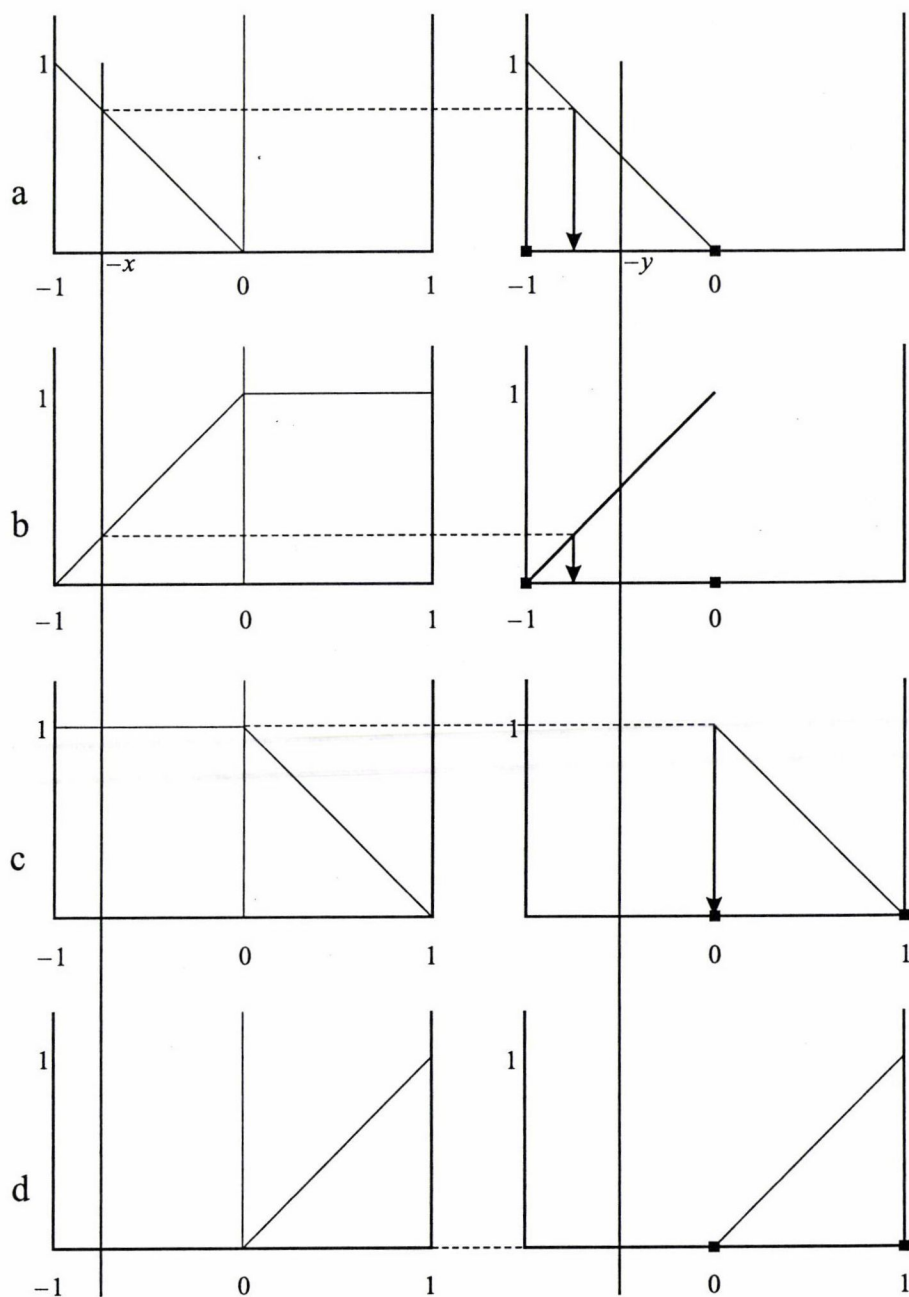
- |  |                                   |   |
|--|-----------------------------------|---|
| a) szabályrendszer: $\mathcal{R}_1$ :    | Ha „ $x$ az <i>pozitiv</i> ”      | <b>akkor</b> „ $y$ az <i>pozitiv</i> ”      |
| b) szabályrendszer: $\mathcal{R}_{21}$ : | Ha „ $x$ az <i>negativ_zero</i> ” | <b>akkor</b> „ $y$ az <i>negativ_zero</i> ” |
| c) szabályrendszer: $\mathcal{R}_{22}$ : | Ha „ $x$ az <i>pozitiv_zero</i> ” | <b>akkor</b> „ $y$ az <i>pozitiv_zero</i> ” |
| d) szabályrendszer: $\mathcal{R}_3$ :    | Ha „ $x$ az <i>negativ</i> ”      | <b>akkor</b> „ $y$ az <i>negativ</i> ”      |

A döntés eredményét a rész döntésekből az SPSG döntéshez hasonlóan számítjuk ki.

Azért, hogy a Tsukamoto döntést az előző eljárásokkal össze lehessen hasonlítani a szabályrendszert úgy állítottuk össze, hogy alakra megegyezzen a kiindulási szabályrendszerrel. Így a döntést leíró kifejezés a következő lett:

$$y_s(x) = \frac{x}{2}.$$

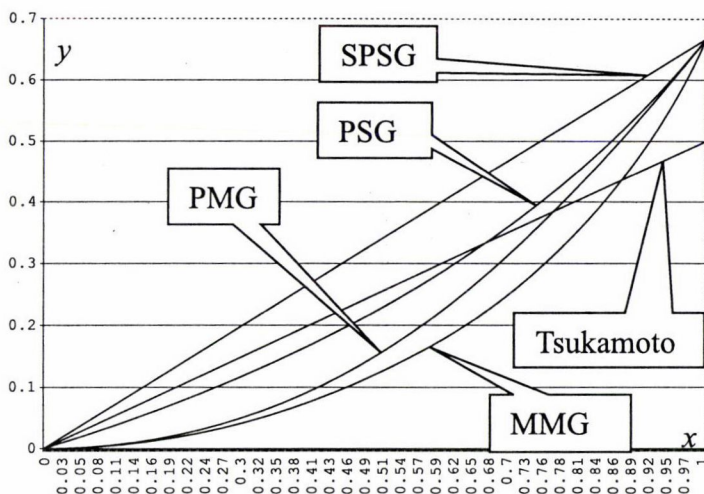




8. ábra: Tsukamoto-féle döntés

#### 4. A döntési eljárások összehasonlítása

A 9. ábra az öt vizsgált döntési eljárás átviteli karakterisztikáját mutatja.



9. ábra: A döntési eljárások karakterisztikái

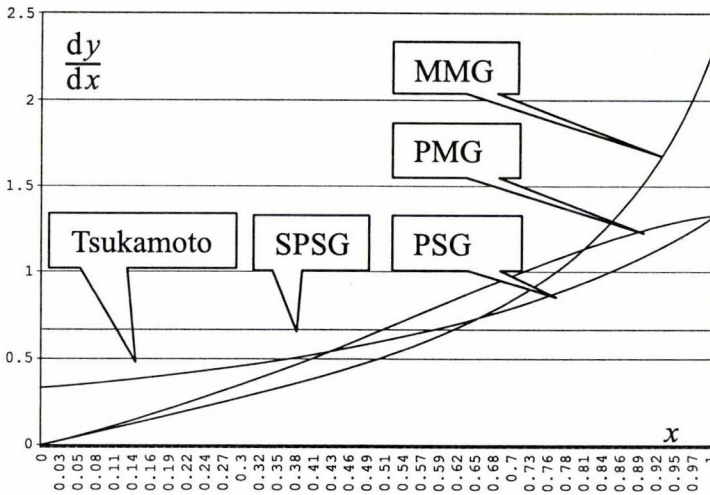
A fenti grafikonok a döntési eljárásoknak csak az  $x \geq 0$  tartományát ábrázolja, ez elegendő mert a függvények páratlanok. A görbékéből láthatjuk, hogy a Tsukamoto-féle döntés kivételével minden döntés  $x = 1$  esetén az  $y = 2/3$  értéket adja. A Tsukamoto-féle döntés az adott esetben  $y = 0,5$  értéket ad. Ezért a további vizsgálatokból a Tsukamoto-féle döntést kihagytuk. Tehát a maximális bementi értékre a döntések azonos kimeneti értéket adnak.

A 10. ábra a döntési karakterisztikák differenciál hányadosát mutatja. A 10. ábrán látható diagrammok alapján láthatjuk, hogy a MMG és a PMG döntések görbéi nullától indulnak, ami azt jelenti, hogy  $x = 0$  környékén a döntések igen „érzékenyek”, a kisebb zavarokat nem viszik át.

A PSG döntés görbéje  $x = 0$  környékén nem nulla, ezért az  $x = 0$  környékén is „érzékeny” a kis eltérésekre. Az SPSG döntés állandó értékű, ez azt jelenti, hogy a döntés a teljes értelmezési tartományon egyformán érzékeny.

A PSG és a PMG döntések az  $x = 1$  helyen (természetesen  $x = -1$  helyen is) azonos értéket mutatnak, ami szerint a két eljárás érzékenysége a végkitérés tartományában azonos.

A MMG döntés görbéje a végkitérés tartományában a legmagasabb, ez azt jelenti, hogy ebben a tartományban a döntés közel kétszer érzékenyebb, mint a PSG és a PMG. A következő táblázat a döntések összehasonlítását foglalja össze:



10. ábra: A döntési karakterisztikák differenciál hányadosa

Döntés	$y_s(x=0)$	$y_s(x=1)$	$\left. \frac{dy_s(x)}{dx} \right _{x=0}$	$\left. \frac{dy_s(x)}{dx} \right _{x=1}$
MMG	0	0,66	0	2,33
PMG	0	0,66	0	1,33
PSG	0	0,66	0,33	1,33
SPSG	0	0,66	0,66	0,66
Tsukamoto	0	0,50	0,50	0,50

1. Tábla: A döntések összehasonlító táblázata

## 5. A döntési eljárások összehasonlítása irányítástechnikai és számítási igény szempontjából

A 10. ábra alapján a MMG és a PMG döntések igen jól használhatók egy típusú szabályozási körökben. Ezek a hurkok tartalmaznak egy integráló típusú jelátvivő tagot. Ez azt jelenti, hogy a stabil állapot a döntés esetén az  $x=0$ . Ekkor a szabályozás nem reagál a kis zavarásokra, viszont nagy eltérések esetén a határozottan beavatkozik, ami ismerve az integráló szabályozási körök „lusta” viselkedését hasznos lehet (nem szabad elfelejtkezni azonban a stabilitási követelményekről sem). Ebből a szempontból a legjobb az MMG döntés. Természetesen ezek a döntések jól alkalmazhatók nulla típusú körökben is főleg akkor, mikor a körerősítés nagy.

A PSG döntés  $x=0$  körüli viselkedése lehetővé teszi, hogy a kis és közepes körerősítésű szabályozási körökben használjuk. A SPSG döntés alkalmazása nem célszerű szabályozási körökben bár alkalmazható, de lineáris jellege (ha más sza-

bályrendszert veszünk fel ez változhat) nem indokolja a fuzzy elv használatát. Ezt a döntést inkább identifikációs problémák megoldására célszerű használni.

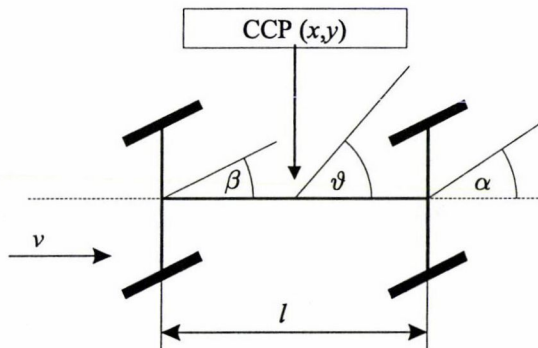
A Tsukamoto döntés szintén lineáris jellegű. Meg kell jegyeznünk, hogy ez a döntés a látszat ellenére megfelelő lenne szabályozó algoritmus kialakítására, de ebben az esetben az összehasonlíthatóság miatt nem térhettünk el a vázolt szabályrendszertől. A fent említett döntések számítási igény szerint alapvetően két csoportra bonthatók.

Az első csoportot az MMG és a PMG alkotják, amelyek a maximum kapcsolatot használják. A számítási igény felbontástól függően változó, de mindenképpen nagy számítási igényű. Ennek oka az, hogy a fuzzy szabályokat célszerűen tömbök segítségével számítják, így egy 3 be – 3 kimentű szabállyal rendelkező döntés 200 elemű szabálysímléír esetén 203 lépést igényel.

A másik csoport az tagjai a PSG és a SPSG, amelyekben a döntések részben előre számíthatók, az előző példát véve ezek algoritmusa 4 lépést, míg Tsukamoto esetben 5 lépést igényelnek.

## 6. Egy gyakorlati példa

A gyakorlati példa egy négykerék kormányzású robottargonca irányításának szimulációja a megfelelő fuzzy döntések segítségével. Az irányítási cél, hogy a robottargoncát a síkban adott  $x, y$  távolságból és  $\vartheta$  állásszög mellett egy adott célterületre irányítsuk. A szimuláció során az MMG, PMG, PSG és az SPSG döntéseket vizsgáljuk. A targonca vázlatát a 11. ábra mutatja.



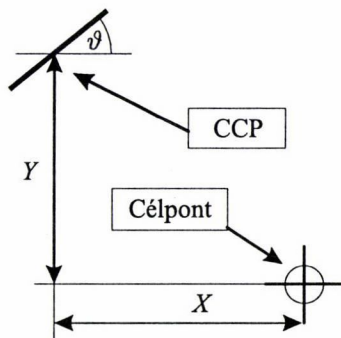
11. ábra: A szimulált kocsí vázlata

A 11-es ábrán a következő jelöléseket alkalmaztuk:  $l$  a kocsí hossza,  $x, y$  a CCP koordinátái,  $\vartheta$  az alváz szöge a világ koordináta-rendszerben,  $v$  a CCP sebessége,  $\alpha$  az első kerekek átlagos szöge a kocsí alvázához képest,  $\beta$  a hátsó kerekek átlagos szöge a kocsí alvázához képest. A szimulációban pedig a következő jellemzőket tételeztük fel:



- a távolságmérő rendszer felbontása 10 mm, a szögmérő rendszer felbontása  $1^\circ$ ,
- a kocsi maximális sebessége 1 m/s, a kocsi minimális sebessége 0,01 m/s, a kormány kitérések felbontása  $1^\circ$ .

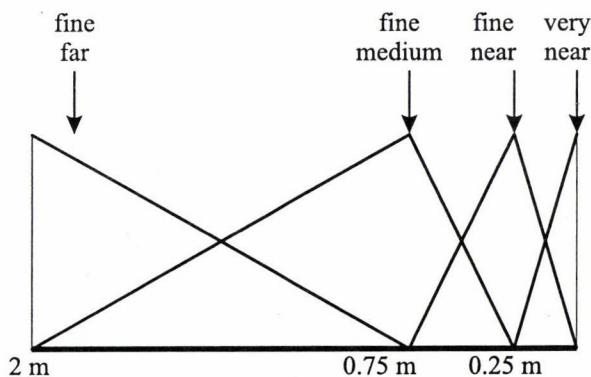
Az irányítási szituációt a 12. ábrán láthatjuk.



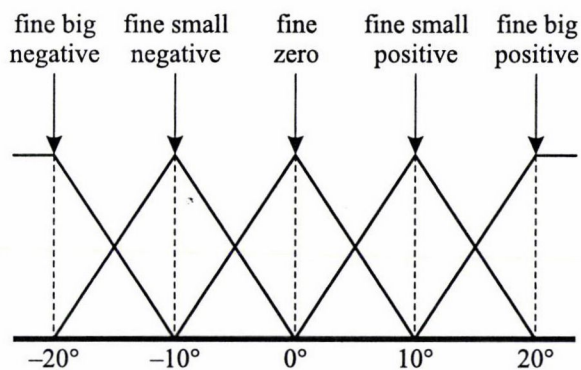
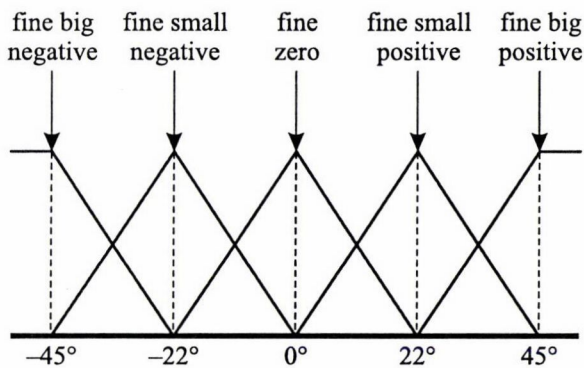
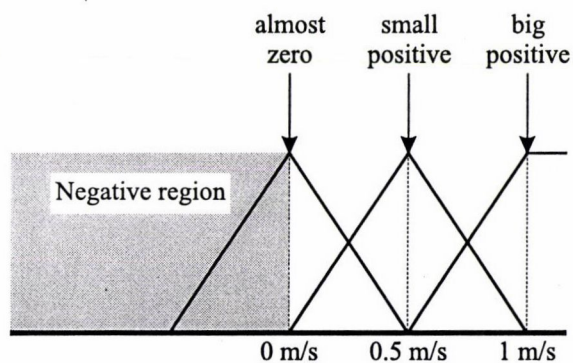
12. ábra: Az irányítási szituáció

A szimuláció a CCP lokációját számítja ki pontról-pontra. Ha tudjuk a kocsi helyzetét egy adott pontban, az adott geometriai jellemzők és a sebesség ismeretével ki tudjuk ezt a pozíciót számítani a következő pontban. Ennek részleteit [4] munkában ismertettük.

A szimuláció két jól elhatárolható szakaszra bontható: egy durva és egy finom megközelítésre. A durva megközelítés feladata, hogy a targoncát a célpont közelébe „elfogadható” jellemzőkkel  $x$ ,  $y$ ,  $\vartheta$  vezesse. A másik részlet a finom megközelítés szakasza, amely a célpont végső megközelítését végzi. Az algoritmusok vizsgálatánál csak ezt a második szakaszt vizsgáljuk, mert ez sokkal „igényesebb” irányítási feladatot jelent. A következő ábrák a bemeneti és a kimeneti szabályrendszereket mutatják:



13. ábra: Bemeneti  $x, y$  szabályrendszerek

14. ábra: Bemeneti  $\vartheta$  szabályrendszer15. ábra: Kimeneti  $\alpha, \beta$  szabályrendszerek16. ábra: Kimeneti  $v$  szabályrendszer

A következő táblázatok az implikációkat meghatározó FAM táblák:

x	F	F	F	F	M	M	M	M	N	N	N	N	VN	VN	VN	VN
$\frac{y}{\vartheta}$	F	M	N	VN	F	M	N	VN	F	M	N	VN	F	M	N	VN
BN	SP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP
SN	SP	SP	BP	BP	SP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP
Z	SN	SN	SN	Z	BN	SN	SN	Z	BN	BN	SN	Z	BN	BN	BN	Z
SP	Z	Z	Z	Z	Z	SN	BN	BN	BN	BN	BN	BN	BN	BN	BN	BN
BP	SN	SN	SN	SN	SN	BN	BN	BN	BN	BN	BN	BN	BN	BN	BN	BN

2. Tábla:  $\alpha$  FAM

x	F	F	F	F	M	M	M	M	N	N	N	N	VN	VN	VN	VN
$\frac{y}{\vartheta}$	F	M	N	VN	F	M	N	VN	F	M	N	VN	F	M	N	VN
BN	SN	SN	SN	SN	SN	BN	BN	BN	BN	BN	BN	BN	BN	BN	BN	BN
SN	Z	Z	Z	Z	Z	SN	BN	BN	BN	BN	BN	BN	BN	BN	BN	BN
Z	SP	SN	SN	Z	BP	SP	SN	Z	BP	BP	SN	Z	BP	BP	BP	Z
SP	SP	SP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP
BP	SP	SP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP

3. Tábla:  $\beta$  FAM

x	F	F	F	F	M	M	M	M	N	N	N	N	VN	VN	VN	VN
$\frac{y}{\vartheta}$	F	M	N	VN	F	M	N	VN	F	M	N	VN	F	M	N	VN
BN	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	SP	SP	SP	SP	SP	Z
SN	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	SP	SP	SP	SP	SP	Z
Z	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	SP	SP	SP	SP	SP	Z
SP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	SP	SP	SP	SP	SP	Z
BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	BP	SP	SP	SP	SP	SP	Z

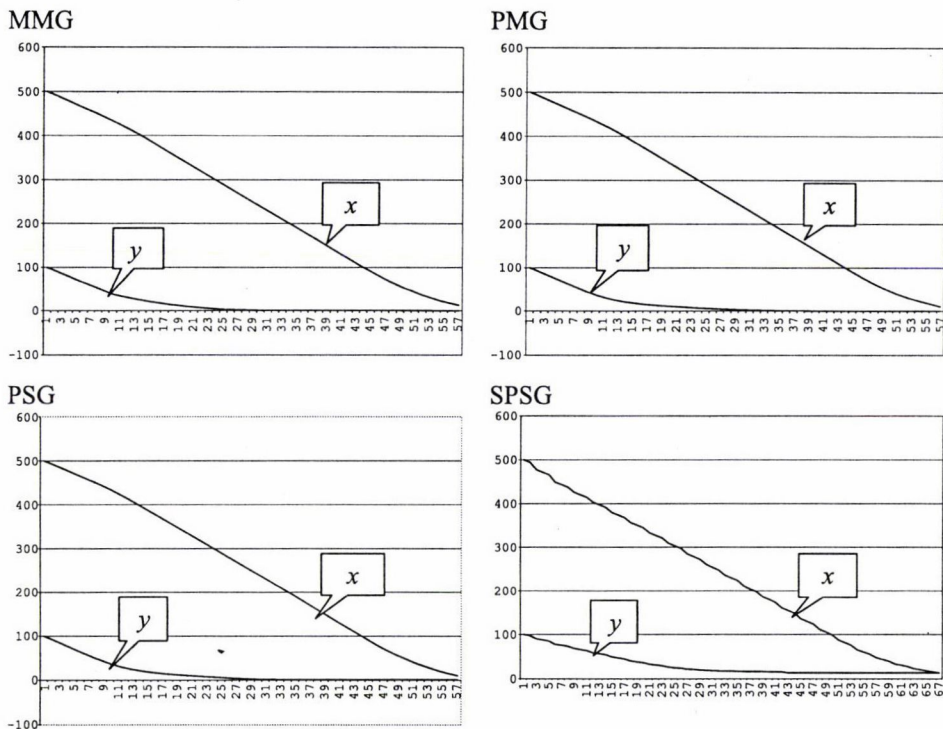
4. Tábla:  $v$  FAM

A különböző döntési eljárások eredményei a szimuláció során a 17. ábrán láthatóak.

A kiindulási értékek a következők voltak:  $x = 5$  m,  $y = 1$  m és  $\vartheta = 0^\circ$ .

## 7. Összefoglalás

A fenti diagrammokból látható, hogy az MMG, a PMG és a PSG döntések gyakorlatilag azonos időben megoldották a feladatot, míg a SPSG reménytelenül eltévedt. látható továbbá, hogy az SPSG a többi döntéshez képest lengéseket végez



17. ábra: A megközelítés  $x$  és  $y$  értékei az idő függvényében

a megközelítés során. A „legsimább” megközelítést az MMG, majd PMG adja, a PSG az előzőekhez képest  $y$  értékben jobban átmegy a negatív tartományba. Az elméleti vizsgálat és a gyakorlati példa eredményei megerősítik egymást. Nem szabad azonban arról a tényről elfeledkezni, ha a szabályrendszerek megváltoznak ez nagyban befolyásolja a fuzzy döntések eredményét. Tipikus példa erre a Tsukamoto-féle döntés, mely az adott gyakorlati feladatra alkalmas lehetett volna, de lényegesen más szabályrendszer és FAM mellett.

## Irodalom

- [1] K. Bonfig, Fuzzy Logik in der industriellen Automatisierung, *Expert Verlag* (1992).
- [2] R. Fullér, Fuzzy Reasoning and Fuzzy Optimization, *TUCS General Publication*, Turku Centre for Computer Science, Åbo, No. 9 (1998).
- [3] M. Mizumoto, Method of fuzzy inference suitable for fuzzy control, *Soc. Instrument Control Engineering*, 58 (1989), 959–963.

- [4] Gy. Schuster, Simulation of fuzzy motion controlled four-wheel steered mobile robot, in: I. J. Rudas ed., *Proceedings of the IEEE International Conference on Intelligent Engineering Systems (INES'97)* (1997), 89–94.
- [5] Y. Tsukamoto, An approach to fuzzy reasoning method, in: M. M. Gupta, R. K. Ragade and R. R. Yager eds., *Advances in Fuzzy Set Theory and Applications* (North-Holland, New York, 1979).
- [6] L. A. ZADEH, Fuzzy Sets, *Information and Control*, 8 (1965), 338–353.
- [7] K. Hirota and L. T. Kóczy, Interpolative rasoning with insufficient evidence in sparse fuzzy rule bases, *Information Science*, 71 (1993), 169–201.
- [8] K. Hirota and L. T. Kóczy, Approximate reasoning by linear rule interpolation and genieral approximation, *International Journal of Approximate Reasoning* (1993), 197–235.
- [9] L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Transactins on Systems, Man and Cybernetics*, 3 (1973), 28–44.

(Beérkezett: 2000. május 30.)

SCHUSTER GYÖRGY  
BUDAPESTI MŰSZAKI FŐISKOLA  
KANDÓ KÁLMÁN VILLAMOSMÉRNÖKI FŐISKOLAI KAR  
MŰSZERTECHNIKAI ÉS AUTOMATIZÁLÁSI INTÉZET  
1084 BUDAPEST  
TAVASZMEZŐ U. 15-17  
E-mail: HAL@K2.JOZSEF.KANDO.HU

## COMPARISON OF FUZZY REASONING METHODS FOR PROCESS CONTROL

GYÖRGY SCHUSTER

We consider five fuzzy decision rules applied widely in practical process control: the minimum-maximum-gravity, the product-maximum-gravity, the product-sum-gravity, the singleton-used-product-sum-gravity and the Tsukamoto reasoning method. We compare these methods not only on their static characteristics and sensitivity, but on the run-time of the controller as well.

After summarizing the necessary definitions and explaining the main aspects of the comparison of fuzzy resoning methods we show a way in which this comparison can be done. Based on the results of the mathematical consideration we classify the reasoning methods connected to control strategies. The paper ends with a practical example.



# BINÁRIS KÓDOK GENERÁLÁSA KÖRNYEZETFÜGGETLEN NYELVTANOK SEGÍTSÉGÉVEL

KÁSZONYI LÁSZLÓ

Szombathely

A [12] cikkben Kászonyi László és Katsura Maszasi bebizonyította a következő állítást: Legyen  $Q$  az  $\{a, b\}$  ábécé feletti primitív szavak nyelve. A  $Q \cap (ab^*)^n$  nyelv környezetfüggetlen, ha az  $n$  szám prímfelbontásában szereplő  $p_1, \dots, p_k$  prímszámokra teljesül, hogy

$$(*) \quad 1/p_1 + \dots + 1/p_k < 4/5$$

Ebben a cikkben megmutatjuk, hogy a  $K_n = Q \cap (a^+b^+)^n$  nyelv is környezetfüggetlen, ha  $(*)$  fennáll. Egyben megadunk véges sok olyan nyelvtant, amelyek együttesen  $K_n$  elemeit generálják.

## 1. Bevezetés

Legyen  $X = \{1, 2\}$  és  $Y = \{a, b\}$  két rögzített ábécé, jelöljük az  $X$  ill.  $Y$  feletti szavak halmazát rendre  $X^*$ -gal, ill.  $Y^*$ -gal.  $X^*$  és  $Y^*$  a szavak egymás után írására mint szorzásra nézve egységelemes szabad félcsoport, azaz szabad monoid. Mindkét félcsoport egységeleme a  $\lambda$ -val jelölt üres szó. ( $\lambda$  olyan szó, amelynek egyetlen betűje sincs.) Legyen  $\phi$   $X^*$ -nak egy művelettartó leképezése  $Y^*$  egy részhalmazára, azaz  $u, v \in X^*$  esetén legyen

$$\phi(uv) = \phi(u)\phi(v),$$

illetve

$$\phi(\lambda) = \lambda.$$

$\phi$ -t *morfizmusnak* mondjuk.  $X^*$  morfizmusai megadhatók oly módon, hogy először megadjuk 0 illetve 1 képét a  $\phi$  leképezésnél, majd a leképezést művelettartó módon kiterjesztjük  $X^*$ -ra. Könnyű belátni, hogy a morfizmusok egy kódolási eljárás algebrai megfelelői: egy  $X$  betűiből készült „szövegben” a 0 ill. 1 betűt rendre a neki megfelelő  $\phi(0)$  ill.  $\phi(1)$  „kódszóra” cseréljük ki. Eljárásunk akkor megfelelő, ha a kódolt szöveg alapján az eredeti szöveget egyértelműen vissza tudjuk állítani,



azaz ha a kód egyértelműen dekódolható. Ennek megfelelően az  $\{u, v\} \subset Y^*$  szópárt akkor és csak akkor nevezzük *kódnak*, ha a  $\phi(0) = u$ ,  $\phi(1) = v$  egyenlőségekkel megadott morfizmus egy-egy értelmű, azaz ha  $\phi$  *izomorfizmus*. Könnyű belátni, hogy  $\{u, v\} \subset Y^*$  akkor és csak akkor kód, ha  $uv \neq vu$  teljesül. (Ld. [18]) A [18] cikkben M. Ito, H. Jürgensen, H. J. Shyr és G. Thierrin vezeti be az  $n$ -kódok fogalmát, itt csak 2-kódokról lesz szó: az  $L \subset Y^*$  szóhalmazt (*nyelvet*) 2-kódnak nevezzük, ha bármely elempárja kód. Míg az  $n$ -kódok jellemzése máig nyílt probléma, a 2-kódok jól leírhatók. A  $p \in Y^+$  szót *primitívnek* nevezzük, ha nem valódi hatvány, azaz ha nem bontható fel valódi módon azonos szakaszokra. (Itt  $Y^+ = Y^* \setminus \{\lambda\}$ .) Könnyű belátni, hogy tetszőleges  $v \in Y^+$  előállítható egy primitív szó hatványaként:  $v = p^n$  ( $n \in \mathbb{N}_+ = \mathbb{N} \setminus \{0\}$ ). Ez az előállítás egyértelmű, a megfelelő  $p$  primitív szót  $v$  *primitív gyökének* mondjuk, és  $\sqrt[n]{v}$ -vel jelöljük. Egyszerűen igazolható, hogy  $u, v \in Y^+$  esetén  $uv \neq vu$  akkor és csak akkor teljesül, ha  $\sqrt[n]{u} \neq \sqrt[n]{v}$ . Ennek alapján egy-egy értelmű megfeleltetést létesítünk a 2-kódok és  $f : Q \rightarrow \mathbb{N}$  típusú függvények között: (ld.: [18]) Legyen  $L \subset Y^+$  tetszőleges 2-kód és adjuk meg az  $f_L : Q \rightarrow \mathbb{N}$  függvényt a következő módon:

$$u \in L \Leftrightarrow f_L(\sqrt[n]{u}) \neq 0 \wedge \sqrt[n]{u}^{f_L(\sqrt[n]{u})} = u.$$

Megfordítva, legyen adva az  $f : Q \rightarrow \mathbb{N}$  függvény. Ennek felhasználásával megadunk egy 2-kódot:

$$L_f = \{p^{f(p)} \mid p \in Q \wedge f(p) \neq 0\}.$$

A 2-kódok fenti reprezentációjából egyrészt levonhatjuk azt az elméleti következtetést, hogy az összes 2-kódok osztálya nem megszámlálható halmaz, így rekurzív felsorolható sem lehet [18], másrészt viszont ötletet nyerünk 2-kódok konstrukciójához. A fentiek szerint ugyanis 2-kódok konstrukciójához elég  $Q$  egy részhalmazát megadnunk, majd a megadott halmaz elemeit hatványoznunk. Mivel köztudott, hogy a veremautomatákkal felismerhető nyelvek osztálya megegyezik a környezetfüggetlen nyelvek osztályával, a 2-kódok vizsgálata  $Q$  részhalmazai környezetfüggetlenségének vizsgálatához vezetett. (Ld.: [4]) Máig eldöntetlen kérdés, hogy a  $Q$  nyelv környezetfüggetlen-e? [1, 2, 3, 9, 10]. Ennek a kérdésnek a vizsgálatával kapcsolatban vetődött fel az  $L_n = Q \cap (ab^*)^n$  nyelv környezetfüggetlenségének kérdése. Sejtésünk az, hogy  $L_n$  minden  $n \in \mathbb{N}$  esetén környezetfüggetlen nyelv. Katsura Maszasival sikerült belátnunk, hogy  $L_n$  valóban környezetfüggetlen, ha az  $n$  szám prímfelbontásában szereplő  $p_1, \dots, p_k$  prímszámokra teljesül, hogy

$$(*) \quad 1/p_1 + \dots + 1/p_k < 4/5$$

Az állítás bizonyítására használt módszer konstruktív abban az értelemben, hogy segítségével 2-kódokat tudunk konstruálni. Ezek a kódok azonban a gyakorlat céljaira nem alkalmasak mivel  $u = ab^{e_0} \dots ab^{e_{n-1}}$  alakúak, azaz  $a$ -ban és  $b$ -ben nem szimmetrikusak. Jobb tulajdonságú kódokat nyerünk a  $K_n = (a^+b^+)^n \cap Q$  halmaz elemeinek generálásával. Ebben a dolgozatban megmutatjuk, hogy a [12] cikkben



megadott módszerrel (\*) fennállása esetén egyrészt belátható  $K_n$  környezetfüggetlensége, másrészt segítségével  $K_n$  elemei generálhatók.

## 2. Alapvető tételek és definíciók

Mint a bevezetőben, legyen  $Y = \{a, b\}$  rögzített ábécé, jelöljük  $Q$ -val az  $Y$  feletti primitív szavak halmazát. A  $G = (N, Y, S, P)$  rendezett négyest *nyelvtannak* mondjuk, ahol  $N$  és  $Y$  két diszjunkt ábécé – a *nemterminálisok* ill. *terminálisok* véges halmaza –,  $S \in N$  az ún. *kezdőszimbólum*, és  $P$  olyan  $(u, v)$  rendezett szópárok véges halmaza, ahol  $u, v \in (N \cup Y)^*$  valamint  $u$  mindig tartalmaz  $N$ -beli elemet. A  $P$ -beli  $(u, v)$  elemeket *produkcióknak* mondjuk, és  $u \rightarrow v$ -vel jelöljük. Azt mondjuk, hogy a  $w'$  szó *egy lépésben levezethető*  $w$ -ből, jelölve:  $w \Rightarrow_1 w'$ , ha létezik olyan  $w_1, u, w_2$  és  $v$  szó, amellyel  $w = w_1 u w_2$ ,  $w' = w_1 v w_2$  és  $u \rightarrow v \in P$  teljesül. A  $w'$  szó  *$k$  lépésben levezethető*  $w$ -ből, jelölve:  $w \Rightarrow_k w'$ , ha van olyan  $N \cup Y$  feletti  $w_0, \dots, w_k$  szósorozat, amelyre teljesül, hogy  $w_0 = w, w_k = w'$  és  $0 \leq i \leq k-1$  esetén  $w_i \Rightarrow_1 w_{i+1}$ . A  $w'$  szó *levezethető* a  $w$  szóból – jelölve:  $w \Rightarrow_* w'$  –, ha van olyan  $k$ , amelyre  $w \Rightarrow_k w'$ . Tetszőleges  $L \subseteq Y^*$  szóhalmazt  $Y$  feletti nyelvnek mondunk. A  $G$  *nyelvtan által generált nyelven* az

$$L(G) = \{w \in Y^* \mid S \Rightarrow_* w\}$$

halmazt értjük. A  $G$  *nyelvtan környezetfüggetlen*, ha a produkciói bal oldalán szereplő szavak egyetlen nemterminális jelből állnak, azaz ha  $u \rightarrow v$  esetén  $u \in N$ . Az  $L$  nyelv *környezetfüggetlen*, ha környezetfüggetlen nyelvtan generálja.

Legyen  $n \in \{1, 2, \dots\}$ , tekintsük  $(a^+ b^+)^n$ -nek egy tetszőleges  $W$  részhalmazát.  $W$ -nek minden  $w$  eleme felírható  $w = a^{e_0} b^{e_1} \dots a^{e_{2n-2}} b^{e_{2n-1}}$  alakban, ahol  $i = 0, \dots, 2n-1$  esetén  $e_i \geq 1$ . Jelöljük a  $w$ -nek megfelelő  $(e_0, \dots, e_{2n-1})$  vektort  $e(w)$ -vel, legyen  $E(W) = \{e(w) \mid w \in W\}$ .

A  $\underline{2n} = \{0, \dots, (2n-1)\}$  indexhalmazt úgy tekintjük, mint „ciklikusan rendezett” halmazt, azaz a  $\underline{2n}$ -beli „nyílt” intervallumokat az  $(i, j) = \{k \mid i < k < j\}$  egyenlőséggel határozzuk meg, ha  $i < j$ , illetve az  $(i, j) = \{k \mid k < j \vee k > i\}$  egyenlőséggel, ha  $i > j$ . A „félig zárt”, ill. „zárt” intervallumot a szokásos módon adjuk meg:  $[i, j) = \{i\} \cup (i, j)$ ,  $(i, j] = (i, j) \cup \{j\}$  és  $[i, j] = \{i\} \cup (i, j) \cup \{j\}$ . A  $\underline{2n}$ -beli összeadást és szorzást úgy tekintjük, mint  $(\text{mod } 2n)$  műveleteket.

Az  $\{i, j\}$  és  $\{k, l\}$  indexpárok *metszők*, ha  $k \in (i, j)$  és  $l \in (j, i)$  vagy ha  $l \in (i, j)$  és  $k \in (j, i)$ . A  $\underline{2n}$  halmazbeli  $R$  és  $T$  halmazok *metszők*, ha létezik olyan  $\underline{2n}$ -beli  $i$  és  $j$ , amelyre teljesül, hogy  $R \subseteq [i, j)$  és  $T \subseteq [j, i)$ . A „nem-metsző” kifejezést  $n.m.$ -mel rövidítjük. Több halmaz esetén a „páronként nem metsző” kifejezés helyett p. n. m.-met írunk.

Az  $L$  nyelvet *korlátosnak* nevezzük, ha léteznek olyan  $w_0, \dots, w_{m-1}$  szavak, amelyekkel  $L$ -nek minden  $w$  eleme felírható  $w = w_0^{e_0} \dots w_{m-1}^{e_{m-1}}$  alakban, ahol  $e_0, \dots, e_{m-1}$  alkalmas természetes számok. Korlátos nyelvek környezetfüggetlenség-

gére vonatkozik Ginsburg [8] tétele, amely szerint az  $L$  korlátos nyelv akkor és csak akkor környezetfüggetlen, ha az

$$E(L) = \{(e_0, \dots, e_{m-1}) \mid w_0^{e_0} \cdots w_{m-1}^{e_{m-1}} \in L\}$$

halmaz rétegesen lineáris halmazok véges egyesítése. (Itt  $w_0, \dots, w_{m-1}$  a korlátos nyelv meghatározásában szereplő szavak.)

Az  $F$  vektorhalmazt *rétegesen lineárisnak* mondjuk, ha  $r \geq 1$ -hez léteznek olyan  $v_0, \dots, v_r \in N^m$  vektorok, amelyekkel

$$F = \left\{ v_0 + \sum_{i=1}^r k_i v_i \mid k_i \geq 0 \right\}$$

és a  $V = \{v_i \mid 1 \leq i \leq r\}$  vektorhalmazra teljesülnek a következők:

- (1) minden  $v \in V$ -nek legfeljebb két nullától különböző komponense van, és
- (2) ha  $u = (u_0, \dots, u_{m-1})$  és  $w = (w_0, \dots, w_{m-1})$  két  $V$ -beli vektor, valamint  $\{i, j\}, \{k, l\}$  metsző indexpárok, akkor  $u_i w_k u_j w_l = 0$ .

Nyilvánvaló, hogy minden  $W \subseteq (a^+ b^+)^n$  nyelv korlátos. Azokat a halmazokat, amelyek rétegesen lineáris halmazok véges egyesítéseként állíthatók elő, *rétegesen szemilineáris* halmazoknak nevezzük. Mivel rétegesen szemilineáris halmazok véges egyesítése újra rétegesen szemilineáris halmaz, ahhoz, hogy belássuk egy  $W$  korlátos nyelv környezetfüggetlenségét, elég ha  $E(W)$ -t előállítjuk véges sok olyan halmaz egyesítéseként, amelyek mindegyike rétegesen szemilineáris halmaz.

Legyen  $n = p_1^{f_1} \cdots p_k^{f_k}$ , ahol  $p_1, \dots, p_k$  páronként különböző prímszámok és  $f_1, \dots, f_k$  pozitív egész számok,  $\pi = \{q_1, \dots, q_r\}$  pedig a  $\{p_1, \dots, p_k\}$  halmaznak egy nem-üres részhalmaza. A  $\{t_1/q_1, \dots, t_r/q_r\}$  halmazt  $\pi$ -skálának nevezzük, ahol  $t_i, i = 1, \dots, r$  esetén  $p_i$ -hez relatív prím szám.

Adott  $S = \{t_1/q_1, \dots, t_r/q_r\}$   $\pi$ -skála és  $\xi \in \underline{2n}$  esetén legyen

$$B = B(\xi; S) = \left\{ \xi + \rho \mid \rho = \sum_{i=1}^r \rho_i t_i 2n/q_i, \rho_i \in \{0, 1\} \right\}.$$

$B$ -t  $\pi$ -doboznak mondjuk. Az  $e = (e_0, \dots, e_{2n-1}) \in \mathbb{N}^{2n}$  vektorhoz tartozó *differencián* értjük a

$$(**) \quad \Delta_e(B) = \Delta_e(\xi; S) = \sum_{\xi + \rho \in B} (-1)^{\rho_1 + \dots + \rho_r} e_{\xi + \rho}$$

formulával megadott kifejezést. Másként szólva, az  $e$  vektorhoz és  $B$  dobozhoz rendelt differencia  $e$  olyan komponenseinek előjeles összege, amelyeknek indexe  $B$ -beli halmaz, és ha az  $(i, j)$  index-pár a  $B$  doboznak „éle”, akkor  $e_i$  és  $e_j$  az összegben ellenkező előjellel szerepelnek.

Főtételünk bizonyításánál szükségünk lesz bizonyos számelméleti fogalmakra, illetve tételekre. Ezek közül itt említjük meg a kínai maradéktételt:

**TÉTEL** (Kinai maradéktétel). Ha  $b_1, \dots, b_k$  tetszőleges egész számok,  $m_1, \dots, m_k$  pedig páronként relatív prímek, akkor mod  $m_1 \cdots m_k$  egyetlen olyan  $x$  egész szám létezik, amelyikre

$$(1) \quad x \equiv b_i \pmod{m_i} \quad (i = 1, \dots, k)$$

teljesül.

### 3. Felhasznált eredmények

Adott  $S$   $\pi$ -skála és  $e \in \mathbb{N}_+^{2n}$  esetén tekintjük  $2n$ -nek az  $\Omega_e(S) = \{\xi \in 2n \mid \Delta_e(\xi; S) \neq \emptyset\}$  részhalmazát. A következőkben azt a kérdést vizsgáljuk, hogy  $\Omega_e(S)$  az üres halmaz-e?

A következő lemma szerint a kérdésre adott válasz független az  $S$  skála választásától.

1. LEMMA. Legyen  $S$  és  $S'$  két  $\pi$ -skála.  $\Omega_e(S) \neq \emptyset$  akkor és csak akkor áll fenn, ha  $\Omega_e(S') \neq \emptyset$ .

*Bizonyítás.* Legyen  $S = \{t_1/q_1, \dots, t_r/q_r\}$  és  $S' = \{t'_1/q_1, \dots, t'_r/q_r\}$ . Bármely  $i$  indexhez létezik olyan  $s_i$  amelyre  $s_i t'_i \equiv t_i \pmod{q_i}$ . Eszerint

$$\Delta_e(\xi; t_1/q_1, \dots, t_r/q_r) = \sum_{j_1=0}^{s_1-1} \dots \sum_{j_r=0}^{s_r-1} \Delta_e(\xi + j_1 t'_1 2n/q_1 + \dots + j_r t'_r 2n/q_r; t'_1/q_1, \dots, t'_r/q_r)$$

Azt mondjuk, hogy  $\Omega_e(\pi) \neq \emptyset$  ha  $\Omega_e(S) \neq \emptyset$  valamely (és így bármely)  $S$   $\pi$ -skála esetén.

A következő lemmában szükséges és elégséges feltételt adunk arra vonatkozóan, hogy egy  $v \in (a^+ b^+)^n$  szó primitív legyen, azaz, hogy  $v \in K_n$  teljesüljön.

2. LEMMA. Legyen  $e = (e_0, \dots, e_{2n-1}) \in \mathbb{N}_+^{2n}$ . Az  $a^{e_0} b^{e_1} \dots a^{e_{2n-2}} b^{e_{2n-1}} \in Q$  reláció akkor és csak akkor áll fenn, ha  $\Omega_e(p_i) \neq \emptyset$  teljesül minden  $i = 1, \dots, k$ -ra.

*Bizonyítás.* Minden  $i = 1, \dots, k$  esetén az  $a^{e_0} b^{e_1} \dots a^{e_{2n-2}} b^{e_{2n-1}}$  szó akkor és csak akkor nem  $p_i$ -hatvány, ha létezik olyan  $j \in 2n$  száma, amelyre teljesül, hogy  $e_j \neq e_{j+2n/p_i}$ .  $\square$

A következő lemma szerint az  $E(K_n)$  vektorhalmaz szemilineáris.

3. LEMMA. Legyen  $K_n = Q \cap (a^+ b^+)^n$ . Ekkor

$$(2) \quad E(K_n) = \cup_{\xi_1, \dots, \xi_k \in 2n} \{(e_0, \dots, e_{2n-1}) \mid e_{\xi_1} \neq e_{\xi_1+2n/p_1}, \dots, e_{\xi_k} \neq e_{\xi_k+2n/p_k}\}.$$

*Bizonyítás.* Vegyük észre, hogy  $u \in K_n$  akkor és csak akkor teljesül, ha  $u$  egyetlen  $i \in \{1, \dots, k\}$  esetén sem  $p_i$  hatvány. Alkalmazzuk az előző lemma állítását  $i = 1, \dots, k$ -ra.  $\square$

Sajnos, mint azt a következő példa mutatja, a (2) felbontás egyes tagjai nem feltétlenül rétegesen szemilineárisak.

1. *Példa.* Legyen  $n = 6 = 2 \cdot 3$ , ekkor  $E(K_6)$ -nak a (2) felbontásában szereplő

$$\{(e_0, \dots, e_{11}) \mid e_0 \neq e_4, e_1 \neq e_7\}$$

tag nem rétegesen szemilineáris. (Ennek bizonyítása pl. az [16]-ben megadott módszerrel lehetséges.)

Másrészről viszont a következő lemma biztosítja, hogy az

$$E(B) = \{e \mid \Delta_e(B) \neq 0\}$$

típusú vektorhalmazok rétegesen szemilineárisak, ezt a tényt fogjuk felhasználni főtételünk bizonyítására.

4. LEMMA (Flip-flop lemma, Kászonyi [14]). Legyen  $i = 0, \dots, m-1$  esetén  $\delta_i$  előjel, azaz legyen  $\delta_i \in \{-1, 0, 1\}$ . Ekkor az

$$E = \{(e_0, \dots, e_{m-1}) \in \mathbb{N}_+^m \mid \delta_0 e_0 + \dots + \delta_{m-1} e_{m-1} \neq 0\}$$

vektorhalmaz rétegesen szemilineáris halmaz.

A  $B_1, \dots, B_v$  dobozok esetén legyen

$$E(B_1, \dots, B_v) = \{e \in \mathbb{N}_+^{2n} \mid \Delta_e(B_1) \neq 0, \dots, \Delta_e(B_v) \neq 0\}.$$

A flip-flop lemma következményeként könnyen bizonyítható a következő lemma állítása:

5. LEMMA. Legyenek  $B_1, \dots, B_v$  páronként nem-metsző dobozok. Ekkor a  $E(B_1, \dots, B_v)$  vektorhalmaz rétegesen szemilineáris.

A következő néhány lemmában a dobozok és a nekik megfelelő differenciák néhány tulajdonságát tárgyaljuk. A 6. Lemma állítása szerint az  $\Omega_e(\pi) \neq \emptyset$  tulajdonság a  $\{p_1, \dots, p_k\}$  halmaz  $\pi$  részhalmazain „öröklődő”.

6. LEMMA. Legyen  $\emptyset \neq \pi' \subseteq \pi \subseteq \{p_1, \dots, p_k\}$ . Ha  $\Omega_e(\pi') = \emptyset$  akkor  $\Omega_e(\pi) = \emptyset$ .

*Bizonyítás.* Legyen  $\pi' = \{q'_1, \dots, q'_{r'}\}$  és  $\pi = \{q'_1, \dots, q'_{r'}, q_1, \dots, q_r\}$ . Ekkor

$$\Delta_e(\xi; 1/q'_1, \dots, 1/q'_{r'}, 1/q_1, \dots, 1/q_r)$$

$$= \sum_{\rho_1=0}^1 \dots \sum_{\rho_r=0}^1 (-1)^{\rho_1+\dots+\rho_r} \Delta_e(\xi + \rho_1 2n/q_1 + \dots + \rho_r 2n/q_r; 1/q'_1, \dots, 1/q'_{r'}). \quad \square$$

7. LEMMA. Legyen  $S$  tetszőleges  $\pi$ -skála és legyen  $q \in \pi$ . Ha  $\xi \in \Omega_e(S)$ , akkor  $\Omega_e(S) \cap \{\xi + j2n/q \mid 1 \leq j \leq q-1\} \neq \emptyset$ .

*Bizonyítás.*  $\sum_{j=0}^{q-1} \Delta_e(\xi + j2n/q; S) = 0$ . □

8. LEMMA. Az  $S$   $\pi$ -skála és  $q \in \{p_1, \dots, p_k\} - \pi$  esetében a következő feltételek ekvivalensek:

- (1)  $\Omega_e(\pi \cup \{q\}) = \emptyset$ .
- (2) Ha  $\xi \equiv \xi' \pmod{2n/q}$  akkor  $\Delta_e(\xi; S) = \Delta_e(\xi'; S)$ .

*Bizonyítás.* Vegyük észre, hogy  $S \cup \{1/q\} \pi \cup \{q\}$ -skála, és hogy  $\Delta_e(\xi; S \cup \{1/q\}) = \Delta_e(\xi; S) - \Delta_e(\xi + 2n/q; S)$  teljesül tetszőleges  $\xi$ -re. Eszerint  $\Omega_e(\pi \cup \{q\}) = \emptyset$ , ha  $\Delta_e(\xi; S) = \Delta_e(\xi + 2n/q; S)$  tetszőleges  $\xi$  esetén. □

9. LEMMA. Az  $S$   $\pi$ -skála és  $\{q_1, \dots, q_r\} \subseteq \{p_1, \dots, p_k\} - \pi$  esetében a következő feltételek ekvivalensek:

- (1)  $\Omega_e(\pi \cup \{q_i\}) = \emptyset$  tetszőleges  $i = 1, \dots, r$ -re.
- (2) Ha  $\xi \equiv \xi' \pmod{2n/q_1 \dots q_r}$  akkor  $\Delta_e(\xi; S) = \Delta_e(\xi'; S)$ .

*Bizonyítás.* Az előző lemma alapján. □

10. LEMMA. Legyen  $S$  tetszőleges  $\pi$ -skála,  $q \in \pi$ , valamint  $\{q_1, \dots, q_r\} \subseteq \{p_1, \dots, p_k\} \setminus \pi$ . Ha  $\xi \in \Omega_e(S)$ , akkor  $\Omega(S) \cap \{\xi + 2jn/q_1 \dots q_r \mid 1 \leq j \leq q-1\} \neq \emptyset$ .

*Bizonyítás.* A 7. és a 9. Lemma alapján. □

11. LEMMA. Tegyük fel, hogy  $i = 1, 2$  esetén  $\Omega_e(\pi_i) \neq \emptyset$  és  $\Omega_e(\pi_i \cup \{q_i\}) = \emptyset$ . Legyen  $S_i$   $\pi_i$ -skála, valamint  $q_1 \neq q_2$ . Ekkor tetszőleges  $\mu \in 2n$ -hez létezik  $\xi_1 \in \Omega_e(S_1)$  és  $\xi_2 \in \Omega_e(S_2) \cap (\xi_1 + \mu, \xi_1 + \mu + 2n/q_1 q_2]$ .

*Bizonyítás.* Legyen  $\rho_1 \in \Omega_e(S_1)$ ,  $\rho_2 \in \Omega_e(S_2)$ , válasszuk meg az  $\alpha, \beta$  számokat oly módon, hogy  $\rho_2 = \alpha 2n/q_1 q_2 + \rho_1 + \mu + \beta$ , és  $\beta \in (0, 2n/q_1 q_2]$  teljesüljön. A Kínai Maradéktétel szerint létezik olyan  $\gamma$  szám, amelyre fennáll  $\gamma \equiv \alpha \pmod{q_1}$  és  $\gamma \equiv 0 \pmod{q_2}$ . Legyen  $\xi_1 = \rho_1 + (\gamma/q_2)(2n/q_1)$  és  $\xi_2 = \xi_1 + \mu + \beta = \rho_2 + ((\gamma - \alpha)/q_1)(2n/q_2)$ . A 8. Lemma szerint  $\xi_1 \in \Omega_e(S_1)$  és  $\xi_2 \in \Omega_e(S_2)$ . □

#### 4. $K_n$ környezetfüggetlensége

Legyen  $n = p_1^{f_1} \dots p_k^{f_k}$ , ahol  $p_1, \dots, p_k$  olyan prímszámok, amelyekre teljesül, hogy  $p_1 < \dots < p_k$ , valamint legyenek  $f_1, \dots, f_k$  tetszőleges pozitív egész számok. Ebben a szakaszban  $K_n = Q \cap (a^+ b^+)^n$  környezetfüggetlenségét bizonyítjuk be bizonyos feltételek teljesülése esetén.

1. TÉTEL. Ha  $(1/p_1 + \dots + 1/p_k) + 1/p_1 p_2 < 1$  és  $(1/p_1 + \dots + 1/p_k) + 1/p_3 < 1$  akkor  $K_n$  környezetfüggetlen nyelv.

(Vegyük észre, hogy  $k = 2$  esetén a második egyenlőtlenség automatikusan teljesül. Hasonlóan, ha  $k = 1$  akkor mindkét egyenlőtlenség igaz.)

A  $\{p_1, \dots, p_k\}$  halmaz  $\pi = \{q_1, \dots, q_r\}$  részhalmaza esetén az  $e \in N^n$  vektorhoz rendelt  $B(\xi; 1/q_1, \dots, 1/q_r)$   $\pi$ -dobozt az egyszerűség kedvéért  $B(\xi; \pi)$ -vel fogjuk jelölni. (Ebben a szakaszban csak ilyen dobozokról lesz szó.)

Legyen  $\{\pi_1, \dots, \pi_v\}$  a  $\{p_1, \dots, p_k\}$  halmaznak egy partíciója, vezessük be az

$$E(\pi_1, \dots, \pi_v) = \cup \{E(B(\xi_1, \pi_1), \dots, B(\xi_v, \pi_v)) \mid B(\xi_1, \pi_1), \dots, B(\xi_v, \pi_v) \text{ p. n. m.}\}$$

jelölést. (Emlékeztetünk a 3. szakaszban bevezetett p. n. m. rövidítésre, amely a „páronként nem metsző” kifejezés helyett áll.)

A 2. szakasz eredményeinek tükrében elég a következőt belátni: Tegyük fel, hogy  $\Omega_e(p_i) \neq \emptyset$  teljesül minden  $i = 1, \dots, k$ -ra. Ekkor létezik a  $\{p_1, \dots, p_k\}$  halmaznak olyan  $\{\pi_1, \dots, \pi_v\}$  partíciója, amelyre  $e \in E(\pi_1, \dots, \pi_v)$  teljesül.

Azt fogjuk belátni, hogy  $\{p_1, \dots, p_k\}$ -nek van olyan  $\{\pi_1, \dots, \pi_v\}$  partíciója, amelyre teljesülnek a következők:

- (1)  $p_i \in \pi_1 \cup \dots \cup \pi_i$  ha  $i = 1, \dots, v$ .
  - (2)  $\Omega_e(\pi_i) \neq \emptyset$  ha  $i = 1, \dots, v$ .
  - (3)  $\Omega_e(\pi \cup \{q\}) = \emptyset$  tetszőleges  $q \in \pi_{i+1} \cup \dots \cup \pi_v$  és  $i = 1, \dots, v-1$  esetén.
  - (4) Ha  $p_i \in \pi_i$  akkor  $\Omega_e(p_{i-1}, p_i) = \emptyset$  ( $i = 2, \dots, v$ ).
- A 10. Lemmából, ill. (3)-ból az is következik, hogy
- (5)  $\Omega_e(\pi_i) \cap (\xi, \xi + 2n / \prod_{q \in \pi_{i+1} \cup \dots \cup \pi_v} q] \neq \emptyset$  bármely  $\xi \in \underline{2n}$  és  $i = 1, \dots, v-1$ -re.

Az (1)–(4) feltételeknek megfelelő partíció adható meg a következő „mohó” algoritmus segítségével:

Tegyük fel, hogy a  $\pi_0, \dots, \pi_i$  halmazok elemeit már kiválasztottuk. (Itt  $\pi_0$  az üres halmaz.) A  $\pi_{i+1}$  halmaz  $q_1, \dots, q_\mu$  elemeit a következő módon választjuk ki:

- Legyen  $q_1$  a  $\{p_1, \dots, p_k\} \setminus \cup_{j=1}^i \pi_j$  halmaz legkisebb eleme.
- Tegyük fel, hogy a  $q_1, \dots, q_s$  elemeket már kiválasztottuk. Ha a következő

$$R_s = \{p_1, \dots, p_k\} \setminus (\cup_{j=1}^i \pi_j) \setminus \{q_1, \dots, q_s\}$$

halmaz üres, akkor legyen  $\mu = s, v = i + 1$  és eljárásunk befejeződik.

Egyébként legyen  $q_{s+1}$   $R_s$ -nek a legkisebb olyan eleme, amelyre teljesül, hogy  $\Omega_e(\{q_1, \dots, q_{s+1}\}) \neq \emptyset$ . Ha nincs ilyen elem, akkor legyen  $\mu = s$ . (Ekkor  $q_s$  a  $\pi_{i+1}$  halmaz utolsó eleme).

Legyen  $\Sigma_i = \sum_{q \in \pi_i} 2n/q$ . Ekkor tételben szereplő feltételek a következők:

$$\Sigma_1 + \dots + \Sigma_v + 2n/p_1 p_2 < 2n, \quad \Sigma_1 + \dots + \Sigma_v + 2n/p_3 < 2n.$$

Be fogjuk látni, hogy  $e \in E(\pi_1, \dots, \pi_v)$ .

A bizonyítás alapötlete a következő: Tegyük fel, hogy van egy, az (1)–(4) feltételeknek megfelelő  $\{\pi_1, \dots, \pi_v\}$  partíciónk. Adott  $\pi \in \{\pi_1, \dots, \pi_v\}$ -hez tekintsünk egy olyan  $\xi_1$ -t amelyre teljesül, hogy  $\Delta_e(B(\xi_1, \pi)) \neq 0$ . Ekkor az előző szakasz eredményeinek felhasználásával meg tudunk adni olyan  $\xi_2 \in \underline{2n}$ -t, amelyre a  $B(\xi_2, \pi)$  doboz megfelelő helyzetű, és amelyre  $\Delta_e(B(\xi_2, \pi)) \neq 0$  is teljesül.

1. eset:  $v = 1$ . (2) alapján  $e \in E(\pi_1)$ .

2. eset:  $v = 2$ ,  $\pi_2 = \{p_2\}$ . (3) és (4) szerint  $\Omega_e(\pi_1 \cup \{p_2\}) = \Omega_e(\pi_2 \cup \{p_1\}) = \emptyset$ . Alkalmazhatjuk tehát a 11. Lemmát, amely szerint létezik  $\xi_2 \in \Omega_e(\pi_2)$  és  $\xi_1 \in \Omega_e(\pi_1) \cap (\xi_2 + \Sigma_2, \xi_2 + \Sigma_2 + 2n/p_1p_2]$ . Mivel

$$B(\xi_2; \pi_2) = \{\xi_2, \xi_2 + \Sigma_2\},$$

$$B(\xi_1; \pi_1) \subseteq [\xi_1, \xi_1 + \Sigma_1] \subseteq (\xi_2 + \Sigma_2, \xi_2 + \Sigma_1 + \Sigma_2 + 2n/p_1p_2] \subseteq (\xi_2 + \Sigma_2, \xi_2 + 2n),$$

$B(\xi_1; \pi_1)$  és  $B(\xi_2; \pi_2)$  nem-metszők, azaz  $e \in E(\pi_1, \pi_2)$ .

3. eset:  $v = 2$ ,  $\pi_2 \neq \{p_2\}$ . Legyen  $\xi_2 \in \Omega_e(\pi_2)$ . Mivel (5) szerint  $\prod_{q \in \pi_2} q \geq p_3$ , létezik  $\xi_1 \in \Omega_e(\pi_1) \cap (\xi_2 + \Sigma_2, \xi_2 + \Sigma_2 + 2n/p_3]$ . Továbbá

$$B(\xi_2; \pi_2) = \{\xi_2, \xi_2 + \Sigma_2\}$$

$$B(\xi_1; \pi_1) \subseteq [\xi_1, \xi_1 + \Sigma_1] \subseteq (\xi_2 + \Sigma_2, \xi_2 + \Sigma_1 + \Sigma_2 + 2n/p_3] \subseteq (\xi_2 + \Sigma_2, \xi_2 + 2n),$$

így  $B(\xi_1; \pi_1)$  és  $B(\xi_2; \pi_2)$  nem-metsző halmazok, azaz  $e \in E(\pi_1, \pi_2)$ .

4. eset:  $v = 3$ , és vagy (a)  $\pi_2 = \{p_i\}$ ,  $\pi_3 = \{p_j\}$  vagy (b)  $|\pi_2| \geq 2$ ,  $\pi_3 = \{p_3\}$ . (3), illetve (4) szerint az (a) esetben  $\Omega_e(\pi_2 \cup \{p_j\}) = \Omega_e(\pi_3 \cup \{p_i\}) = \emptyset$  és a (b) esetben  $\Omega_e(\pi_2 \cup \{p_3\}) = \Omega_e(\pi_3 \cup \{p_2\}) = \emptyset$ .

Mivel  $2n/p_i p_j \leq 2n/p_2 p_3$ , a 11. Lemma szerint létezik  $\xi_3 \in \Omega_e(\pi_3)$  és  $\xi_2 \in \Omega_e(\pi_2) \cap (\xi_3 + \Sigma_3, \xi_3 + \Sigma_3 + n/p_2 p_3]$ . (5) szerint  $\prod_{q \in \pi_2 \cup \pi_3} q \geq p_2 p_3$ , azaz

$$\xi_1 \in \Omega_e(\pi_1) \cap (\xi_3 + \Sigma_2 + \Sigma_3 + 2n/p_2 p_3, \xi_3 + \Sigma_2 + \Sigma_3 + 2n/p_2 p_3].$$

Mivel

$$B(\xi_3; \pi_3) = \{\xi_3, \xi_3 + \Sigma_3\},$$

$$B(\xi_2; \pi_2) \subseteq [\xi_2, \xi_2 + \Sigma_2] \subseteq (\xi_3 + \Sigma_3, \xi_3 + \Sigma_2 + \Sigma_3 + 2n/p_2 p_3],$$

$$\begin{aligned} B(\xi_1; \pi_1) &\subseteq [\xi_1, \xi_1 + \Sigma_1] \subseteq (\xi_3 + \Sigma_2 + \Sigma_3 + 2n/p_2 p_3, \xi_3 + \Sigma_1 + \Sigma_2 + \Sigma_3 + 2n/p_2 p_3] \\ &\subseteq (\xi_3 + \Sigma_2 + \Sigma_3 + 2n/p_2 p_3, \xi_3 + \Sigma_1 + \Sigma_2 + \Sigma_3 + 2n/p_3] \\ &\subseteq (\xi_3 + \Sigma_2 + \Sigma_3 + 2n/p_2 p_3, \xi_3 + 2n), \end{aligned}$$

nyerjük, hogy  $e \in E(\pi_1, \pi_2, \pi_3)$ .

5. eset:  $v = 3$ , de sem (a) sem (b) nem teljesül. Tekintsük  $\xi_3 \in \Omega_e(\pi_3)$ -t. Mivel  $\prod_{q \in \pi_3} q \geq p_4$  és  $\prod_{q \in \pi_2 \cup \pi_3} q \geq p_3 p_4$ , létezik

$$\xi_2 \in \Omega_e(\pi_2) \cap (\xi_3 + \Sigma_3, \xi_3 + \Sigma_3 + 2n/p_4],$$

$$\xi_1 \in \Omega_e(\pi_1) \cap (\xi_3 + \Sigma_2 + \Sigma_3 + 2n/p_4, \xi_3 + \Sigma_1 + \Sigma_2 + \Sigma_3 + 2n/p_4 + 2n/p_3 p_4].$$

Mivel  $2n/p_4 + 2n/p_3p_4 < 2n/p_3$ , a 4. esethez hasonlóan nyerjük, hogy  $e \in E(\pi_1, \pi_2, \pi_3)$ .

6. eset:  $v \geq 4$ . Legyen  $\xi_v \in \Omega_e(\pi_v)$ . Mivel  $\prod_{q \in \pi_v} q \geq p_v$  és  $\prod_{q \in \pi_{i+1} \cup \dots \cup \pi_v} q \geq p_{v-1}p_v$ , létezik

$$\begin{aligned} \xi_{v-1} &\in \Omega_e(\pi_{v-1}) \cap (\xi_v + \Sigma_v, \xi_v + \Sigma_v + 2n/p_v], \\ \xi_{v-2} &\in \Omega_e(\pi_{v-2}) \cap (\xi_v + \Sigma_v + \Sigma_{v-1} + 2n/p_v, \xi_v + \Sigma_v + \Sigma_{v-1} + 2n/p_v + 2n/p_v p_{v-1}] \\ \xi_{v-3} &\in \Omega_e(\pi_{v-3}) \cap (\xi_v + \Sigma_v + \Sigma_{v-1} + \Sigma_{v-2} + 2n/p_v + 2n/p_v p_{v-1}, \\ &\quad \xi_v + \Sigma_v + \Sigma_{v-1} + \Sigma_{v-2} + 2n/p_v + 2n/p_v p_{v-1}] \\ &\quad \vdots \\ \xi_{v-j} &\in \Omega_e(\pi_{v-j}) \cap (\xi_v + \Sigma_v + \dots + \Sigma_{v-(j-1)} + 2n/p_v + (j-2)2n/p_v p_{v-1}, \\ &\quad \xi_v + \Sigma_v + \dots + \Sigma_{v-(j-1)} + 2n/p_v + (j-1)2n/p_v p_{v-1}] \end{aligned}$$

ahol  $j \in \{2, \dots, v-1\}$ .

Mivel

$$\begin{aligned} 1/p_v + (v-2)/p_{v-1}p_v &\leq 1/p_3 - (p_v - p_3)/p_3p_v + (v-2)/p_{v-1}p_v \\ &\leq 1/p_3 - 2(v-3)/p_3p_v + (v-2)/p_{v-1}p_v \leq 1/p_3 - (v-4)/p_3p_v \leq 1/p_3, \end{aligned}$$

a  $B(\xi_i; \pi_i)$  ( $i = 1, \dots, v$ ) boxok páronként nem-metszők, és így  $e \in E(\pi_1, \dots, \pi_v)$ . Könnyű belátni, hogy

$$(3) \quad E(K_n) = \cup \{E(\pi_1, \dots, \pi_v) \mid \{\pi_1, \dots, \pi_v\} \in \Pi\},$$

ahol  $\Pi$  a  $\{p_1, \dots, p_k\}$  halmaz particióinak összessége.

$E(K_n)$  rétegesen szemilineáris halmazok véges egyesítése, azaz maga is rétegesen szemilineáris, így Ginsburg tétele szerint  $K_n$  környezetfüggetlen.  $\square$

1. Következmény. Ha  $1/p_1 + \dots + 1/p_k < 4/5$ , akkor  $K_n$  környezetfüggetlen.

2. Következmény. Ha  $1/p_1 + \dots + 1/p_k < 4/5$ , akkor

$$(4) \quad E(K_n) = \bigcup_{\{\pi_1, \dots, \pi_v\} \in \Pi} E(\pi_1, \dots, \pi_v) = \bigcup_{\{\pi_1, \dots, \pi_v\} \in \Pi} \cup \{E(B(\xi_1, \pi_1), \dots, B(\xi_v, \pi_v)) \mid \xi_1, \dots, \xi_v \in \underline{2n}, B(\xi_1, \pi_1), \dots, B(\xi_v, \pi_v) \text{ p. n. m.}\}$$

Itt is, mint az 1. Tétel bizonyításánál,  $\Pi$ -vel a  $\{p_1, \dots, p_k\}$  halmaz összes  $\{\pi_1, \dots, \pi_v\}$  alakú particióját jelöltük.



## 5. Kódok konstrukciója

Ebben a szakaszban megmutatjuk, hogy hogyan lehet az előző szakasz eredményeit felhasználni  $K_n$  elemeinek generálására. Adott  $n \in \mathbb{N}$  esetén olyan  $G$  környezetfüggetlen nyelvtanokat adunk meg, amelyek  $K_n$ -beli elemeket generálnak. A konstrukcióhoz felhasználjuk az ún. flip-flop tételt, amely a flip-flop lemma (4. Lemma) általánosítása. A tétel kimondásához szükségünk lesz a következő definícióra:

1. *Definíció.* Az

$$(5) \quad E(\Theta, \delta, \varepsilon, R) = \bigcap_{I \in \Theta} \left\{ (e_0, \dots, e_{m-1}) \in \mathbb{N}^m \mid \varepsilon(I) \sum_{i \in I} \delta_i e_i R(I) 0 \right\}$$

halmazt **DLI**-halmaznak mondjuk, ahol

(1)  $\Theta$  indexhalmazok nem-üres rendszere, (azaz  $\underline{m}$  részhalmazzaiból álló halmazrendszer).

(2)  $\delta = (\delta_0, \dots, \delta_{m-1})$  rögzített *előjelvektor*, azaz  $i = 0, \dots, m-1$  esetén  $\delta_i \in \{-1, 0, +1\}$ .

(3)  $\varepsilon$  olyan  $\Theta$ -n értelmezett függvény, amelynek értékkészlete a  $\{+1, -1\}$  halmaz.

(4) Az  $R(I)$  vagy „ $\geq$ ”-t vagy „ $>$ ”-et jelöl.

Ha ez nem okoz félreértést, akkor az  $E(\Theta, \delta, \varepsilon, R)$  jelölésből elhagyjuk  $R$ -et azaz röviden  $E(\Theta, \delta, \varepsilon)$ -t írunk.

2. *Definíció.* Az  $L$  korlátos nyelvet **DLI-nyelvnek** mondjuk, ha az  $L$  nyelvhez rendelt exponensvektorok

$$E(L) = \{(e_0, \dots, e_{m-1}) \in \mathbb{N}^m \mid w_0^{e_0} \cdots w_{m-1}^{e_{m-1}} \in L\}$$

halmaza **DLI**-halmaz.

A  $K_n$  nyelv nyilvánvalóan **DLI**-nyelv akár (2), akár (4) alapján.

A **DLI**-nyelvek környezetfüggetlenségének vizsgálata Ginsburg tétele szerint egyenértékű a megfelelő **DLI**-halmazok réteges szemilinearitásának vizsgálatával. A következő tétel az alapja azoknak a vizsgálatoknak, amelyeknek célja bizonyos **DLI**-halmazok réteges szemilinearitásának megállapítása.

**TÉTEL** (flip-flop tétel) [16]. Legyen  $E$  a  $\Theta$  indexhalmaz-rendszer, a  $\delta$  előjelvektor és  $\varepsilon$  függvény által meghatározott **DLI**-halmaz.  $E$  akkor és csak akkor rétegesen szemilineáris, ha minden  $e \in E$ -hez létezik olyan  $H$  hipergráf, amelyre teljesülnek a következő feltételek:

(i)  $H$  pontjai egy konvex  $m$ -szög csúcsai, amelyeket ciklikus sorrendjük szerint a  $0, \dots, m-1$  számokkal azonosítunk.

(ii)  $H$  élei  $H$  pontjaiból álló egy- vagy kételemű halmazok.

(iii) Ha  $\{i, j\}$   $H$ -nak kételemű éle, akkor az  $i$  és  $j$  pontokhoz rendelt előjelekre  $\delta_i = -\delta_j \neq 0$  teljesül.

(iv) Az  $f$  él tiltott, ha létezik olyan  $I \in \Theta$  indexhalmaz, amelyre  $f \cap I = \{i\}$  és  $\varepsilon(I) = -\delta_i$  teljesül.  $H$  nem tartalmaz tiltott éleket.

(v) Az  $f = \{i, j\}$  és  $g = \{k, l\}$  élek esetén tegyük fel, hogy  $i < j$  és  $k < l$ .  $f$  és  $g$  metsző élek, ha vagy  $i < k < j < l$  vagy  $k < i < l < j$  teljesül.  $H$  nem tartalmaz metsző éleket.

(vi) Ha  $e = (e_0, \dots, e_{m-1})$ , akkor az  $i$  pont  $H$ -beli foka  $e_i$ . ( $i = 0, \dots, m-1$ ).

A tétel [16]-ban adott bizonyítása konstruktív abban az értelemben, hogy ha ismerjük azt a  $H$  hipergráfot, amelyet a flip-flop tétel szerint egy  $e \in E$  vektorhoz rendelünk hozzá, akkor meg tudjuk határozni  $E$ -nek azt az  $E_e$  rétegesen szemilineáris komponensét, amelynek  $e$  eleme:

– Először megadjuk azt a  $H'$  hipergráfot, amelynek ugyanazok a pontjai és élei, mint  $H$ -nak, de  $H'$ -ben minden szereplő él egyszeres multiplicitású.

– A  $H'$  hipergráf felhasználásával az  $E_e$  szemilineáris halmaz  $P_e$  periódushalmazát a következő módon határozzuk meg:

$$(6) \quad P_e = \{v(h) \mid h \in E(H'), v(h) = (v_0, \dots, v_{m-1}), \\ v_i = 1, \text{ ha } i \in h, v_i = 0 \text{ ha } i \notin h\}$$

ahol  $E(H')$  a  $H'$  gráf élhalmazát jelöli.

A bizonyításból az is kiderül, hogy  $H'$  speciálisan választható meg, elérhető, hogy  $H'$  komponensei fagráfok legyenek. (Vegyük észre, hogy a  $H'$  csak annyiban tér el a közönséges gráfoktól, hogy tartalmazhat egy pontú éleket is.  $H'$  komponenseit akkor mondjuk fagráfnak, ha kétpontú éleik fagráftól feszítenek ki.)

A következőkben rátérünk a  $K_n$ -net generáló nyelvtanok konstrukciójának vázlatos ismertetésére. Tekintsük  $E(K_n)$ -nek a (4)-ben adott felbontását. Az itt szereplő komponensek egyes rétegesen lineáris összetevőikhez fogunk egy-egy nyelvtant hozzárendelni.

A  $G$  hipergráfot a  $\Theta$  indexhalmazra és  $\delta$  előjelvektorra nézve *rétegesnek* nevezzük, ha eleget tesz a flip-flop tételben szereplő (i)–(v) feltételeknek. A következőkben olyan műveletet vezetünk be, amely réteges gráfhoz réteges gráfot rendel. Tekintsük  $G$ -nek egy tetszőleges  $i$  pontját, jelöljük az  $i$ -hez illeszkedő egy-pontú, ill. kétpontú élek számát rendre  $r_i^1$ -gyel, ill.  $r_i^2$ -vel. („Egyfok”, ill. „kettőfok”). Adott  $m^1$  és  $m^2$  ( $0 \leq m^1 \leq r_i^1$ ,  $0 \leq m^2 \leq r_i^2$ ) esetén képezzük a következő,  $\kappa(G) = \kappa(G, i, m^1, m^2)$ -vel jelölt réteges gráfot:

– Hasítsuk ketté  $G$ -nek az  $i$  pontját, azaz hagyjuk el  $G$  pontjai közül  $i$ -t és pótoljuk az  $i_1$ , ill.  $i_2$  pontokkal.  $i_1$  és  $i_2$  ciklikus sorrend szerinti elhelyezése legyen  $\dots i-1, i_1, i_2, i+1 \dots$ .

– Tekintsük  $G$ -nek az  $i$  ponthoz illeszkedő kétpontú éleiből álló  $E_i^2 = \{h_1, \dots, h_{v_i^2}\}$  élhalmazt. Tegyük fel, hogy  $E_i^2$  elemeinek indexezése megfelel a ciklikus sorrendnek, azaz ha  $h_j = \{i, s\}$ ,  $h_k = \{i, t\}$  és  $j < k$ , akkor  $s \leq t$  is teljesül. (Itt a

„<” relációjel ciklikus sorrendet jelöl, azaz  $j < j + k$  ha  $k = 1, \dots, m - 1$ .  $G$ -nek lehetnek többszörös élei is, így előfordulhat, hogy  $s = t$ .)

Hasítsuk ketté az  $E_i^2$  élhalmazt: a  $h_1, \dots, h_{m^2}$  élek  $i$ -től különböző végpontját kössük össze  $i_2$ -vel, a  $h_{m^2+1}, \dots, h_{v^2}$  pontokét pedig  $i_1$ -gyel.

– Hasítsuk ketté az  $i$  ponthoz illeszkedő egy pontú éleket is, azaz vegyük hozzá  $\kappa(G)$ -hez az  $\{i_2\}$  élnek  $m^1$  db, az  $\{i_1\}$  élnek pedig  $r_i^1 - m^1$  db. példányát.

– Legyen  $\delta_\kappa(i_1) = \delta_\kappa(i_2) = \delta(i)$  ill.  $\delta_\kappa(j) = \delta(j)$ , ha  $j \neq i$ . A  $\Theta_\kappa$ -val jelölt halmazrendszer úgy nyerjük a  $\Theta$  halmazrendszerből, hogy minden olyan halmazban, amelyben szerepel az  $i$  elem, ezt két elemmel,  $i_1$ -gyel, ill.  $i_2$ -vel helyettesítjük. Könnyű belátni, hogy  $\kappa(G)$  a  $\delta_\kappa$  előjel függvényre, ill. a  $\Theta_\kappa$  indexhalmaz-rendszerre vonatkozóan réteges hipergráf.

– Végül számozzuk újra  $\kappa(G)$  pontjait a  $0, \dots, i - 1, i_1, i_2, i + 1, \dots, m - 1$  sorrendnek megfelelően. (Azaz  $\kappa(G)$  ponthalmaza legyen a  $\{0, \dots, m\}$  halmaz.) A következőkben magát a  $G \rightarrow \kappa(G)$  hozzárendelést is  $\kappa$ -val fogjuk jelölni.

A  $\kappa$  operáció  $\kappa^{-1}$ -gyel jelölt megfordítását olyan gráf homomorfizmusnak foghatjuk fel, amely a  $\kappa(G)$  gráf két szomszédos, azonos  $\delta$  előjelű,  $i_1$ -gyel ill.  $i_2$ -vel jelölt pontját az  $i$  pontra húzza össze, a gráf többi pontját, ill. élet pedig helyben hagyja.

A  $G$  réteges hipergráfot *eleminek* mondjuk, ha pontjainak fokszáma legfeljebb egy. (Esetünkben olyan elemi gráfok fognak szerepelni, amelyeknek minden pontja elsőfokú.) Nyilvánvaló, hogy tetszőleges  $G$  réteges hipergráf – esetleg több lépésben – széthasítható elemi gráffá, azaz létezik hasításoknak olyan  $\kappa_1, \dots, \kappa_\rho$  sorozata, amelynek elemeit egymás után alkalmazva elemi gráfot nyerünk. Legyen  $\kappa = \kappa_\rho \circ \dots \circ \kappa_1$ , ahol a  $\circ$  műveleti jel operációk kompozícióját jelöli. Nyilvánvaló, hogy  $\kappa^{-1} = \kappa_1^{-1} \circ \dots \circ \kappa_\rho^{-1}$  gráf homomorfizmusnak is felfogható. Legyen

$$E = \left\{ v + \sum_{e \in P} k_e e \mid k_e \geq 0 \right\},$$

$E(K_n)$ -nek egy, a (4)-ben szereplő rétegesen lineáris komponense,  $P$  periódushalmazzal és  $v$  preperiódussal. Jelöljük  $H'$ -vel a  $P$ -hez rendelt réteges hipergráfot,  $H_\kappa = \kappa(H')$ -vel egy olyan elemi réteges hipergráfot, amelyet  $H'$  széthasításával nyerünk.

A  $H_\kappa$  és a  $H'$  hipergráf, ill. a  $v = (v_0, \dots, v_{m-1})$  preperiódus felhasználásával határozzuk meg a  $G = (N, Y, S, P)$  nyelvtant a következő módon:

– Jelöljük  $m$ -mel a  $H_\kappa$  hipergráf pontszámát,  $E = E^1 \cup E^2$ -vel élhalmazát, ahol  $E^1$ , ill.  $E^2$  rendre az egy pontú ill. a két pontú élek halmaza.

– Legyen  $N = \{S\} \cup \{U_j \mid j \in \underline{m}\} \cup \{Z_{ij} \mid 0 \leq i < j \leq m - 1\} \cup \{V_h \mid h \in E(H_\kappa)\}$ ,

–  $Y = \{a, b\}$

– A  $P$  produkcióhalmaz algoritmikus leírása a következő:

(a) Először  $S$  helyére  $Z_{0,m-1}$ -et helyettesítünk:  $S \rightarrow Z_{0,m-1}$ .

- (b) Mivel  $H_\kappa$  minden pontja elsőfokú, van a 0 ponthoz illeszkedő  $h \in \mathbf{E}(H_\kappa)$  éle. Ha  $h = \{0\} \in \mathbf{E}^1$ , akkor helyettesítsünk  $Z_{0,m-1}$  helyére

$$U_0^{v_0} V_{\{0\}} Z_{1,m-1}\text{-et, ha pedig } h = \{0, j\} \in \mathbf{E}^2, \text{ akkor}$$

$$U_0^{v_0} V_{\{0,j\}} U_j^{v_j} Z_{j+1,m-1}\text{-et } (j = m \text{ esetén legyen } Z_{j+1,m-1} = \lambda.) :$$

$$Z_{0,m-1} \rightarrow U_0^{v_0} V_{\{0\}} Z_{1,m-1}, \quad \text{ha } h = \{0\} \in \mathbf{E}^1$$

$$Z_{0,m-1} \rightarrow U_0^{v_0} V_{\{0,j\}} U_j^{v_j} Z_{j+1,m-1}, \quad \text{ha } h = \{0, j\} \in \mathbf{E}^2 \text{ és } j < m-1.$$

$$Z_{0,m-1} \rightarrow U_0^{v_0} V_{\{0,j\}} U_j^{v_j}, \quad \text{ha } h = \{0, j\} \in \mathbf{E}^2 \text{ és } j = m-1.$$

- (c) Hasonlóan járhatunk el általában a  $Z_{ij}$  nemterminális esetében. Legyen  $h$   $H_\kappa$ -nak az  $i$  ponthoz illeszkedő éle. Ha  $h = \{i\} \in \mathbf{E}^1$ , akkor helyettesítsünk  $Z_{ij}$  helyére  $U_i^{v_i} V_{\{i\}} Z_{i+1,j}$ -t, ha pedig  $h = \{i, k\} \in \mathbf{E}^2$ , akkor  $U_i^{v_i} V_{\{i,k\}} U_k^{v_k} Z_{k+1,j}$ -t. Az utóbbi helyettesítést úgy értjük, hogy amennyiben  $j = k$ , akkor  $Z_{k+1,j} = \lambda$ :

$$Z_{i,j} \rightarrow U_i^{v_i} V_{\{i\}} Z_{i+1,j}, \quad \text{ha } h = \{i\} \in \mathbf{E}^1$$

$$Z_{i,j} \rightarrow U_i^{v_i} V_{\{i,k\}} U_k^{v_k} Z_{k+1,j}, \quad \text{ha } h = \{i, k\} \in \mathbf{E}^2, \text{ és } k < j$$

$$Z_{i,j} \rightarrow U_i^{v_i} V_{\{i,k\}} U_k^{v_k}, \quad \text{ha } h = \{i, k\} \in \mathbf{E}^2, \text{ és } k = j.$$

- (d) A  $h$  gráfélhez rendelt  $V_h$  nemterminális fő szerepe az, hogy segítségével „iterációs sorozatot” képezzünk: Ha  $h = \{i\}$ , akkor tekintsük

$$V_{\{i\}} \rightarrow U_i V_{\{i\}}\text{-t, ha } h = \{i, j\}, \text{ akkor pedig } V_{\{i,j\}} \rightarrow U_i V_{\{i,j\}} U_j\text{-t.}$$

- (e) Az  $U_i$ -k iterálását a következő módokon zárhatjuk le:

$$V_{\{i\}} \rightarrow \lambda$$

$$V_{\{i,j\}} \rightarrow Z_{i+1,j-1}, \quad \text{ha } (i, j) \neq \emptyset \quad (\text{Azaz ha } j > i+1)$$

$$V_{\{i,j\}} \rightarrow \lambda, \quad \text{ha } (i, j) = \emptyset \quad (\text{Azaz ha } j = i+1).$$

- (f) Végül kialakítjuk az előállított szó végső formáját, az  $U_i$ -k helyére terminálisokat helyettesítünk:

$$U_i \rightarrow \begin{cases} a, & \text{ha } \kappa^{-1}(i) \text{ páratlan,} \\ b, & \text{ha } \kappa^{-1}(i) \text{ páros.} \end{cases}$$

(Emlékeztetünk arra, hogy a  $\kappa^{-1}$  leképezés mint gráfhomomorfizmus a  $H_\kappa$ , hipergráf egy-egy pontjához  $H$ -beli pontot, azaz a  $2n$  halmaz egy elemét rendeli hozzá.)

*Megjegyzések.* Az a kérdés, hogy a  $K_n$  nyelv környezetfüggetlen-e olyan  $n$ -nek esetében is, amelyekre (\*) nem teljesül, továbbra is nyitott marad. Az általunk adott bizonyításban a (\*) feltételnek döntő szerepe van: biztosítja, hogy a  $B_1, \dots, B_v$  dobozok elhelyezhetők egy  $2n$  hosszúságú szakaszon anélkül, hogy metszenék egymást. Több kísérletet tettünk a cikkben használt módszer továbbfejlesztésére, ezek azonban nem vezettek eredményre. További eredményeket várunk a [12] és [13] cikkekben megadott módszerektől.

## References

- [1] P. Dömösi, S. Horváth and M. Ito, Formal languages and primitive words, *Publ. Math. (Debrecen)*, **42** (1993), 315–321.
- [2] P. Dömösi, S. Horváth, M. Ito, L. Kászonyi and M. Katsura, Some combinatorial properties of words and the Chomsky-hierarchy, *Second International Colloquium on Words, Languages and Combinatorics*, Kyoto, August 25–28, 1992. In: *Proc. 2nd Colloquium on Words, Languages and Combinatorics*, Kyoto, 1992. Ed. by M. Ito and H. Jürgensen, World Scientific (Singapore, 1994), 105–123.
- [3] P. Dömösi, S. Horváth, M. Ito, L. Kászonyi and M. Katsura, Formal languages consisting of primitive words, *Fundamentals of Computer Science* (Szeged, 1993), *Lecture Notes in Computer Science*, Springer-Verlag, **710** (1993), pp. 194–203.
- [4] M. Ito and M. Katsura, Context-free languages consisting of non-primitive words. *Intern. J. Computer Math.*, **40** (1991), 157–167.
- [5] Gh. Păun, Morphisms and Primitivity, *EATCS Bulletin, Probleme Column* (1995).
- [6] A. Salomaa, From Parikh vectors to GO territories, *EATCS Bulletin, Formal Language Column* (1995).
- [7] J. Berstel and L. Boasson, The set of Lyndon words is not context-free *EATCS Bulletin, Probleme Column* (1996).
- [8] S. Ginsburg, *The Mathematical Theory of Context-Free Languages*, McGraw-Hill Book Comp. (New York, 1966).
- [9] H. Petersen, The ambiguity of primitive words, in: *Proc. STACS'94*, Springer-Verlag (Berlin, etc., 1994).
- [10] A. Mateescu, Gh. Păun, G. Rosenberg and A. Salomaa, Parikh prime words and GO-like territories, *Journal of Universal Computer Science*, **1**(12) (1995), 790–810.
- [11] L. Kászonyi and M. Katsura, On the context-freeness of a class of primitive word, *Publicationes Math.* (Debrecen, 1997).
- [12] L. Kászonyi and M. Katsura, Some new results on the context – freeness of languages  $Q \cap (ab^*)^n$ , *AFL'97, Salgótarján, Hungary*, in: *Publicationes Mathematicae* (Debrecen, 1998).
- [13] L. Kászonyi and M. Katsura, On an algorithm concerning the languages  $Q \cap (ab^*)^n$ , *P.U.M.A. (Pure and Applied Mathematics.)*
- [14] L. Kászonyi, On a class of stratified linear sets, *First joint conference on modern applied mathematics*, Ilieni/Ilye (Romania, 1995), in: *P.U.M.A.*, **6** (1995), No. 2, pp. 203–210.
- [15] L. Kászonyi and M. Holzer, On the generalized Flip-Flop Lemma, *AFL'97, Salgótarján, Hungary*, in: *Publicationes Mathematicae* (Debrecen, 1998).

- [16] L. Kászonyi, On bounded context-free languages, in: T. Imaoka and C. Nehaniv, editors, *Proceedings of the First Symposium on Algebra, Languages and Computation*, University of Aizu (1997), Japan.
- [17] L. Kászonyi, Hypergraps and stratified semilinear DLI-sets, Extended abstract for the Japanese-Hungarian Symposium on Discrete Mathematics and its Applications, (Kyoto, 1999), in: *Proceedings of the conference*.
- [18] M. Ito, H. Jürgensen, H. J. Shyr and G. Thierrin, Languages whose  $n$ -element subsets are codes, *Theoretical Computer Science*, **96** (1992), 325–344.

(Beérkezett: 2000. június 9.)

KÁSZONYI LÁSZLÓ  
 BERZSENYI DÁNIEL FŐISKOLA  
 MATEMATIKA TANSZÉK  
 H-9700 SZOMBATHELY  
 KÁROLYI G. TÉR 4  
 E-mail: KASZONYI@FS2.BDTF.HU

## HOW TO GENERATE BINARY CODES USING CONTEXT-FREE GRAMMARS

LÁSZLÓ KÁSZONYI

It is proved by László Kászonyi and Masashi Katsura in [12] that if  $Q$  is the language of primitive words defined over the alphabet  $\{a, b\}$  then the language  $Q \cap (ab^*)^n$  is context-free whenever for prime components  $p_1, \dots, p_k$  of the natural number  $n$  holds

$$(*) \quad 1/p_1 + \dots + 1/p_k < 4/5$$

In this paper it is shown that assuming  $(*)$ , the language  $K_n = Q \cap (a^+b^+)^n$  is context-free as well.

## ERGODIKUS MARKOV-LÁNCOK ÉS ALKALMAZÁSAIK

FARAGÓ ISTVÁN ÉS KOVÁCS MIHÁLY

Budapest

A Markov-folyamatok és a Markov-láncok a természettudomány számos területén alapvető szerepet játszanak. A Markov-láncok vizsgálatának egyik alapproblémája a rendszer viselkedésének vizsgálata hosszú idő eltelte után: megfigyelhető-e a rendszer egyfajta „stabilizálódása” egy jól meghatározott állapot körül? Dolgozatunkban a skáláris ergodikus Markov-láncok egy egységes, az egylépéses iterációs módszereken alapuló felépítését mutatjuk be, emellett alapvető fizikai alkalmazások is ismertetésre kerülnek.

### 1. Bevezetés

A Markov-folyamatok és a Markov-láncok a természettudomány számos területén játszanak alapvető szerepet ([5], [9]). Elég ha arra gondolunk, hogy a Markov-folyamatok a fizikából ismert kauzalitási elv matematikai megfelelői, azaz olyan folyamatokat modelleznek, melyeknél a rendszer jövőbeli viselkedését csak a jelenlegi állapot határozza meg. (A folyamatnak nincsen „emlékezete”.). A Markov-láncok olyan speciális Markov-folyamatok, ahol az idő diszkrét lépésekben telik. A Markov-láncok vizsgálatának egyik alapproblémája a rendszer viselkedésének vizsgálata hosszú idő eltelte után: megfigyelhető-e a rendszer egyfajta „stabilizálódása” egy jól meghatározott állapot körül ún. (stacionárius állapot), avagy ilyen nem történik. Az ergodicitás fogalma a folyamat stacionárius állapotához kapcsolódik, így ennek vizsgálata kiemelt fontosságú ([7], [8]).

A Markov-láncok matematikai vizsgálata során három különböző megközelítéssel találkozhatunk. Az első a valószínűségelméleti megközelítés, amely a Markov-láncokat speciális sztochasztikus folyamatként kezeli. A második, a lineáris algebrai vizsgálati mód, amely elsősorban az elemi osztók elméletére épül. A harmadik megközelítési mód a gráfelméleten alapul, ahol speciális tulajdonságú gráfosztálynak tekintjük a Markov-láncokat. Az egyes tárgyalásmódok saját eszköztáruk felhasználásával építik fel a Markov-láncok elméletét. Ebből eredendően bizonyos eredmények vagy csak az egyes területeken jelennek meg, vagy lényegében ugyan-

azon eredmények különböző formában, más terminológiával szerepelnek. (Példaként megemlítjük az ergodikusság fogalmának kettősségét.)

Dolgozatunk célja a következő:

1. a különböző megközelítések közötti kapcsolat feltárása;
2. egy egységes, az egy lépéses iterációs módszereken alapuló, az egyes részek eredményeit felhasználó tárgyalásmód felépítése;
3. fontos fizikai alkalmazások bemutatása.

Az irodalom az egyes megközelítéseket részletesen tárgyalja. A lineáris algebrai megközelítéshez az [3], [4], [10], [12] munkák szolgálhatnak alapul. A valószínűségelméleti megközelítés az [5], [7], [8], míg a gráfelméleti tárgyalásmód az [1], [9] könyvekben szerepel.

A dolgozat felépítése a következő. A második fejezetben összefoglaljuk azon lineáris algebrai tételeket és definíciókat, amelyeket a Markov-láncok vizsgálatánál használunk. A bevezetett, lényeges fogalmakat fizikai háttérrel rendelkező példákkal illusztráljuk. A harmadik szakasz a diszkrét paraméterű véges állapotterű Markov-folyamatokról szól. Az alapvető fogalmak mellett példákat adunk Markov-láncokra, alkalmazva az előző szakasz példáit. A negyedik és az ötödik fejezetben az ergodikus Markov-láncokat tárgyaljuk. Megkülönböztetve az eloszlás-ergodikus és az osztály-ergodikus láncokat, megadjuk az eloszlás-ergodicitás feltételeit. Az osztály-ergodikus Markov-láncokra összefoglaljuk az ismert eredményeket. Befejezésül a hatodik szakaszban fizikai példákön szemléltetjük az eloszlás- illetve osztály-ergodikus Markov-láncokat.

Dolgozatunkban több új állítás, megjegyzés illetve néhány jól ismert tétel az általunk ismerttől eltérő és lényegesen egyszerűbb bizonyítása szerepel. A többnyire közismert lineáris algebrai tételek illetve a dolgozat felépítése szempontjából nem lényeges tételek bizonyításai megtalálhatók a hivatkozott irodalomban ([1], [3], [4], [10], [12]).

Dolgozatunk eredményei a skaláris Markov-láncokra vonatkoznak, emellett lehetőséget teremtenek a többdimenziós általánosítás irányába is. Ugyanakkor nem foglalkozunk a Markov-láncok numerikus kérdéseivel. Az erre vonatkozó eredmények részben megtalálhatók a [2], [6], [11] cikkekben.

## 2. Lineáris algebrai alapok

Először összefoglaljuk a dolgozat tárgyalásához szükséges lineáris algebrai ismereteket.

**2.1. Definíció.** Azt mondjuk, hogy egy  $A = \{a_{ij}\} \in \mathbb{R}^{N \times N}$  mátrix reducibilis, ha létezik olyan  $P$  permutáló mátrix, hogy az  $E = PAP^T$  mátrix

$$E = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$



alakú, ahol  $\mathbf{B}$  és  $\mathbf{D}$  négyzetes mátrixok, (vagy  $n = 1$  estén  $\mathbf{A} = 0$ ). Ha  $\mathbf{A}$  nem reducibilis, akkor irreducibilisnek nevezzük.

**2.2. Definíció.** Azt mondjuk, hogy egy  $\mathbf{A} \in \mathbb{R}^{N \times N}$  mátrix nemnegatív (pozitív), ha elemeire az  $a_{ij} \geq 0$ , ( $a_{ij} > 0$ )  $i, j = 1, 2, \dots, N$  feltétel teljesül.

**Jelölés:**  $\mathbf{A} \geq 0$ , ( $\mathbf{A} > 0$ ).

Nyilvánvalóan a fenti definíció a vektorokra, mint speciális mátrixokra ugyanígy értelmezhető.

**2.1. ÁLLÍTÁS (Frobenius).** Legyen  $\mathbf{A} \in \mathbb{R}^{N \times N}$  nemnegatív irreducibilis mátrix, és jelölje  $\rho(\mathbf{A})$  a mátrix spektrálsugarát. Ekkor

- $\rho(\mathbf{A})$  pozitív;
- $\rho(\mathbf{A})$  az  $\mathbf{A}$  mátrix sajátértéke;
- $\rho(\mathbf{A})$  egyszeres sajátérték;
- létezik olyan  $\mathbf{w} \in \mathbb{R}^N$  pozitív sajátvektor, amelyre  $\mathbf{A}\mathbf{w} = \rho(\mathbf{A})\mathbf{w}$ .

**2.2. ÁLLÍTÁS.** Ha  $\mathbf{A} \in \mathbb{R}^{N \times N}$  nemnegatív mátrix, akkor

$$\min_j \left\{ \sum_{i=1}^N a_{ij} \right\} \leq \rho(\mathbf{A}) \leq \max_j \left\{ \sum_{i=1}^N a_{ij} \right\}$$

illetve

$$\min_i \left\{ \sum_{j=1}^N a_{ij} \right\} \leq \rho(\mathbf{A}) \leq \max_i \left\{ \sum_{j=1}^N a_{ij} \right\}.$$

**2.3. Definíció.** Egy  $\mathbf{A} \in \mathbb{R}^{N \times N}$  nemnegatív irreducibilis mátrixot  $h$  indexű imprimitív mátrixnak nevezünk, ha  $h$  számú  $\rho(\mathbf{A})$  abszolút értékű sajátértéke van. Ha  $h = 1$  akkor a mátrixot primitívnek nevezzük.

**2.3. ÁLLÍTÁS.** Egy  $\mathbf{A} \in \mathbb{R}^{N \times N}$  nemnegatív mátrix pontosan akkor primitív, ha létezik olyan  $m \in \mathbb{N}$  szám, amelyre  $\mathbf{A}^m > 0$ .

**2.4. ÁLLÍTÁS.** Egy  $\mathbf{A} \in \mathbb{R}^{N \times N}$  irreducibilis mátrix primitív, ha  $\text{tr } \mathbf{A} > 0$ , ahol  $\text{tr } \mathbf{A} = \sum_{i=1}^N a_{ii}$ .

A továbbiakban a dolgozat szempontjából fontos gráfelméleti ismereteket foglalkoztatjuk össze.

**2.4. Definíció.** Egy  $\mathbf{A} \in \mathbb{R}^{N \times N}$  mátrixhoz rendelt irányított gráfnak nevezzük azt a  $P_1, P_2, \dots, P_N$  csúcsokat tartalmazó irányított gráfot, amelyben  $P_i$ -ből  $P_j$ -be pontosan akkor vezet él, amikor  $a_{ij} \neq 0$ .

**Jelölés:**  $G(\mathbf{A})$ .

**2.5. Definíció.** Azt mondjuk, hogy egy  $G$  irányított gráf erősen összefüggő, ha tetszőleges, a gráf csúcsaiból alkotott  $(P_i, P_j)$  rendezett pár esetén létezik a gráf éleinek olyan sorozata amely  $P_i$ -ből  $P_j$ -be vezet. Ezt a  $P_i$ -ből  $P_j$ -be vezető

útnak nevezzük. Legyen  $R = \{P_{i_0}P_{i_1}, P_{i_1}P_{i_2}, \dots, P_{i_{l-1}}P_{i_l}\}$  egy út a  $G$  gráfban. Az  $l$  számot az  $R$  út hosszának nevezzük. Azt mondjuk, hogy  $R$  kör ha  $P_{i_0} = P_{i_l}$ .

2.5. ÁLLÍTÁS. Egy  $\mathbf{A} \in \mathbb{R}^{N \times N}$  mátrix pontosan akkor irreducibilis, amikor  $G(\mathbf{A})$  erősen összefüggő.

Imprimitív mátrixok esetén a mátrix irányított gráfjának vizsgálatával kiszámítható a mátrix imprimitivitási indexe.

2.6. ÁLLÍTÁS. Legyen  $\mathbf{A} \in \mathbb{R}^{N \times N}$  egy irreducibilis mátrix. Jelölje  $S_i$  az összes  $G(\mathbf{A})$ -beli  $P_i$ -t tartalmazó, körök hosszainak halmazát, továbbá  $h_i$  az  $S_i$ -beli számok legnagyobb közös osztóját. Ekkor  $h_1 = h_2 = \dots = h_N = h$ , és  $\mathbf{A}$   $h$  indexű imprimitív mátrix.

2.1. Példa. Tekintsük az alábbi  $\mathbb{R}^{N \times N}$ -es mátrixokat:

$$\mathbf{A}_1 = \begin{pmatrix} 0 & 1 & 0 & . & . & . & 0 \\ q & 0 & p & . & . & . & 0 \\ 0 & q & 0 & p & 0 & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & q & 0 & p \\ 0 & . & . & . & 0 & 1 & 0 \end{pmatrix};$$

$$\mathbf{A}_2 = \begin{pmatrix} q & p & 0 & . & . & . & 0 \\ q & 0 & p & . & . & . & 0 \\ 0 & q & 0 & p & 0 & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & q & 0 & p \\ 0 & . & . & . & 0 & q & p \end{pmatrix};$$

$$\mathbf{A}_3 = \begin{pmatrix} 0 & p & 0 & . & . & . & q \\ q & 0 & p & . & . & . & 0 \\ 0 & q & 0 & p & 0 & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & q & 0 & p \\ p & . & . & . & 0 & q & 0 \end{pmatrix},$$

ahol  $0 < p < 1$  és  $q = 1 - p$ . Nyilvánvalóan 2.5. Állítás alapján mindhárom mátrix irreducibilis mivel a hozzájuk rendelt irányított gráfok erősen összefüggőek. A mátrixokhoz rendelt irányított gráfokról leolvasható, hogy  $\mathbf{A}_1$  2 indexű imprimitív mátrix,  $\mathbf{A}_2$  primitív mátrix,  $\mathbf{A}_3$ , páros  $N$  esetén 2 indexű imprimitív mátrix, míg páratlan  $N$  esetén primitív mátrix. A fenti példákból is jól látszik, hogy  $\text{tr } \mathbf{A} > 0$  csak egy elégséges feltétele a primitívségnek. Az  $\mathbf{A}_1$  mátrix irányított gráfja  $N = 4$  esetén az 1. ábrán látható, az  $\mathbf{A}_2$  mátrix irányított gráfját  $N = 4$  esetén a 2. ábra

mutatja és az  $A_3$  mátrix irányított gráfja  $N = 5$  illetve  $N = 6$  esetén a 3. ábrán illetve a 4. ábrán látható.

**2.6. Definíció.** Azt mondjuk, hogy egy  $A \in \mathbb{R}^{N \times N}$  mátrix szemikonvergens, ha létezik a  $\lim_{j \rightarrow \infty} A^j$  határérték.

Tekintsük az

$$(1) \quad x^{(k+1)} = Ax^{(k)}, \quad k = 0, 1, \dots, \quad x^{(0)} \text{ adott}$$

iterációt.

**2.7. Definíció.** Azt mondjuk, hogy az (1) iteráció szemikonvergens a  $H \subset \mathbb{R}^N$  halmazon, ha az (1) iteráció által generált  $x^{(k)}$  vektor sorozat minden  $x^{(0)} \in H$  esetén konvergens. Ha az iteráció határvektora nem függ a kezdeti vektor megválasztásától, akkor az iterációt konvergensnek nevezzük a  $H$  halmazon. (Ha  $H = \mathbb{R}^N$ , akkor az iterációt szemikonvergensnek, illetve konvergensnek nevezzük.)

Az  $A$  mátrix Jordan alakjából nyilvánvaló az alábbi állítás.

**2.7. ÁLLÍTÁS.** Az (1) iteráció pontosan akkor szemikonvergens, ha  $A$  szemikonvergens.

**Megjegyzés.** Ha egy iteráció szemikonvergens, akkor az  $x$ ,  $x^{(0)}$ -tól függő határvektor fixpontja az  $A : \mathbb{R}^N \rightarrow \mathbb{R}^N$  leképezésnek, másrészt a leképezés tetszőleges fixpontja előállítható az (1) iterációval egy megfelelő kezdővektorból kiindulva. Ezenkívül az iteráció határértékben minden kezdeti vektort az  $Ax = x$  egyenlet egy megoldásába viszi át és az egyenlet tetszőleges megoldása előállítható az (1) iterációval egy megfelelő kezdővektorból kiindulva.

**Megjegyzés.** Az (1) iteráció pontosan akkor szemikonvergens, ha szemikonvergens az  $\mathbb{R}^N$  egy olyan részhalmazán, amely tartalmazza az  $\mathbb{R}^N$  tér egy bázisát.

**2.8. Definíció.** Azt mondjuk, hogy egy  $A = \{a_{ij}\} \in \mathbb{R}^{N \times N}$  mátrix sztochasztikus mátrix, ha  $A$  nemnegatív és  $\sum_{j=1}^N a_{ij} = 1$ ,  $i = 1, \dots, N$ .

**2.1. Következmény.** **2.2. Állítás** alapján, ha  $A$  sztochasztikus mátrix, akkor  $\rho(A) = 1$ .

**2.8. ÁLLÍTÁS.** Egy  $A$  sztochasztikus mátrix pontosan akkor szemikonvergens, ha a  $\lambda = 1$  az egyetlen egységnyi sajátértéke.

### 3. Diszkrét paraméterű véges állapotterű Markov-folyamatok

A Markov-láncok a sztochasztikus folyamatok speciális esetei. Általában a sztochasztikus folyamatokat a következőképpen definiálhatjuk. Legyen  $I \subset \mathbb{R}$  egy halmaz,  $\{X_t\}_{t \in I}$  valószínűségi változók egy családja. Sztochasztikus folyamatnak nevezzük az  $(\{X_t\}_{t \in I}, I)$  párt. Ha  $I$  megszámlálható, akkor diszkrét paraméterű sztochasztikus folyamatnak nevezzük, ha az  $\{X_t\}_{t \in I}$  valószínűségi változók értékkészlete véges, akkor pedig véges állapotterű sztochasztikus folyamatról beszélünk. Ha egy sztochasztikus folyamat diszkrét paraméterű és véges állapotterű, akkor Markov-láncnak nevezzük. A nem diszkrét paraméterű Markov-folyamatokat mértékelméleti alapon definiáljuk, így azok vizsgálata a lineáris algebrától eltérő eszközöket igényel. A nem véges állapotterű Markov-folyamatok átmenetvalószínűségek mátrixa végtelen dimenziós, ezért ezekre az itt kimondott tulajdonságok nem mind érvényesek. A nem homogén-Markov-láncok esetén az átmenetvalószínűségek mátrixa  $k$ -tól is függ (azaz attól, hogy hányadik lépésről van szó) és ez az eset speciális vizsgálatot igényel ([5], [7], [8]).

Tekintsünk egy olyan rendszert, amely csak  $N$  különböző állapotban lehet. Jelölje  $S = \{s_1, \dots, s_N\}$  a rendszer lehetséges állapotait tartalmazó véges halmazt, továbbá  $\mathbf{T} = \{t_{ij}\} \in \mathbb{R}^{N \times N}$  azt a mátrixot, ahol  $t_{ij}$  annak a valószínűsége, hogy a rendszer az  $s_i$  állapotból az  $s_j$  állapotba kerül;  $\mathbf{p}^{(0)} = (p_1^{(0)}, \dots, p_N^{(0)})$  pedig jelentse azt a vektort, ahol  $p_i^{(0)}$  annak a valószínűsége, hogy a rendszer kezdetben az  $s_i$  állapotban tartózkodik. Nyilván  $p_i^{(0)} \geq 0$  minden  $i = 1, \dots, N$  esetén és  $\sum_{i=1}^N p_i^{(0)} = 1$ .

**3.1. Definíció.** A  $(\mathbf{T}, \mathbf{p}^{(0)}, S)$  rendezett hármast diszkrét paraméterű véges állapotterű homogén Markov-folyamatnak, röviden Markov-láncnak nevezzük.

Megjegyezzük, hogy a  $(\mathbf{T}, \mathbf{p}^{(0)}, S)$  hármásban az  $S$  halmazt szokásos még állapotterűnek, a  $\mathbf{T}$  mátrixot az átmenetvalószínűségek mátrixának, a  $\mathbf{p}^{(0)}$  vektort pedig kezdeti valószínűség-eloszlásnak is nevezni.

**3.1. ÁLLÍTÁS.** Legyen  $(\mathbf{T}, \mathbf{p}^{(0)}, S)$  egy Markov-lánc. Ekkor  $\mathbf{T}$  sztochasztikus mátrix.

*Bizonyítás.* Ha a rendszer valamely  $s_i$  állapotban van, akkor annak a valószínűsége, hogy a rendszer az  $s_1, \dots, s_N$  állapotok valamelyikébe kerül, eggyel egyenlő. Mivel ezek egymást kizáró események, ezért az események összegének valószínűsége megegyezik az események valószínűségének összegével, ami pontosan  $\sum_{j=1}^N t_{ij}$ , azaz  $\mathbf{T}$  sztochasztikus mátrix.

Jelölje  $\mathbf{p}^{(n)} = (p_1^{(n)}, \dots, p_N^{(n)})$  azt a valószínűségeloszlás vektort, ahol  $p_i^{(n)}$  annak a valószínűsége, hogy a rendszer  $n$  lépés után az  $s_i$  állapotban van. A továbbiakban a  $\mathbf{p}^{(n)}$  vektort az  $n$ -edik valószínűség-eloszlás vektornak fogjuk nevezni.

Nyilvánvalóan tetszőleges  $i = 1, \dots, N$  és  $n = 0, 1, \dots$  esetén  $p_i^{(n)} \geq 0$  és  $\sum_{i=1}^N p_i^{(n)} = 1$ .

*Megjegyzés.* A továbbiakban a  $\mathbf{p}$  vektort sorvektornak tekintjük, az egyszerűség kedvéért ezt külön nem jelöljük.

3.1. LEMMA [1]. Legyen  $(\mathbf{T}, \mathbf{p}^{(0)}, S)$  egy Markov-lánc. Ekkor minden  $n$  esetén:

$$\mathbf{p}^{(n)} = \mathbf{p}^{(n-1)}\mathbf{T} = \mathbf{p}^{(0)}\mathbf{T}^n.$$

3.1. Példa (Markov-lánca). Tekintsünk egy részecskét amely egységnyi lépésekben egy egyenesen mozog. Tegyük fel, hogy  $p$  pozitív valószínűséggel lép jobbra, és  $q = 1 - p$  pozitív valószínűséggel lép balra, továbbá a részecske csak véges sok  $s_i$  helyre léphet két határoló pont  $s_1$  és  $s_N$  között. Feltesszük továbbá, hogy ha a részecske eléri a határoló pontokat, akkor 1 valószínűséggel visszalép. Ezt a rendszert modellezhetjük egy Markov-lánccal, melynek  $\mathbf{T}$  átmenetvalószínűségek mátrixa a 2. fejezetben szereplő  $\mathbf{A}_1$  típusú, azaz  $\mathbf{T} = \mathbf{A}_1$ . Ezt a láncot egydimenziós korlátos véletlen bolyongásnak nevezzük és számos fizikai jelenség modellezésére alkalmazzák (például diffúzió). A  $\mathbf{T} = \mathbf{A}_2$  és a  $\mathbf{T} = \mathbf{A}_3$  eset (ez utóbbit egydimenziós korlátlan bolyongásnak is nevezik) is ugyanezt a folyamatot írja le csupán a peremfeltételek különböznek, és ez mint láttuk jelentősen befolyásolja a mátrixok primitívségét.

#### 4. Eloszlás-ergodikus Markov-lánccok

4.1. Definíció. Egy  $(\mathbf{T}, \mathbf{p}^{(0)}, S)$  Markov-láncot eloszlás-ergodikusnak nevezzük, ha  $\mathbf{p}^{(0)}$  megválasztásától függetlenül léteznek a

$$\lim_{n \rightarrow \infty} \mathbf{p}^{(n)} = \mathbf{p}^*$$

határértékek.

4.1. ÁLLÍTÁS. Egy  $(\mathbf{T}, \mathbf{p}^{(0)}, S)$  Markov-lánc pontosan akkor eloszlás-ergodikus, ha  $i$ -től függetlenül léteznek a

$$\lim_{n \rightarrow \infty} T_{ik}^n = T_k$$

határértékek, ahol  $T_{ik}^n = \{\mathbf{T}^n\}_{ik}$

*Megjegyzés.* Vegyük észre, hogy itt valójában két feltétel is van:

- $\mathbf{T}$  szemikonvergens;
- a határmátrix speciális szerkezetű, nevezetesen, minden sora ugyanaz.

*Bizonyítás.* Tegyük fel, hogy  $i$ -től függetlenül léteznek a

$$\lim_{n \rightarrow \infty} T_{ik}^n = T_k$$

határértékek. Mivel

$$p_k^{(n)} = \sum_{i=1}^N p_i^{(0)} T_{ik}^n,$$

ezért tetszőleges  $\mathbf{p}^{(0)}$  esetén:

$$\lim_{n \rightarrow \infty} p_k^{(n)} = \sum_{i=1}^N p_i^{(0)} T_k = T_k.$$

Most tegyük fel, hogy  $\mathbf{p}^{(0)}$  megválasztásától függetlenül létezik a

$$\lim_{n \rightarrow \infty} \mathbf{p}^{(n)} = \mathbf{p}^*$$

határérték. Így 2.7. Állítás alapján  $\mathbf{T}$  szemikonvergens mátrix, azaz létezik  $\lim_{k \rightarrow \infty} \mathbf{T}^k = \mathbf{T}^*$ . Ezért

$$\mathbf{p}^* = \mathbf{p}^{(0)} \mathbf{T}^*$$

tetszőleges  $\mathbf{p}^{(0)}$  esetén. Végül  $\mathbf{p}^{(0)}$ -nak választva rendre az  $\mathbb{R}^N$  szokásos bázisvektorait, az állítást kapjuk.

*Megjegyzés.* A fenti bizonyítás során nem használtuk fel azt, hogy a  $\mathbf{p}^{(0)}$  vektor elemei nemnegatívak. Ezért, ha a Markov-lánc definíciójából kihagynánk a nemnegativitási feltételt, a fenti tétel továbbra is érvényes maradna. (Ekkor azonban nem tudunk neki valószínűségi értelmezést adni, mivel a valószínűség-eloszlástól megköveteljük a nemnegativitást.) Vegyük észre, hogy a tétel még általánosabb formában is érvényes. Legyen ugyanis

$$\mathbf{p}^{(0)} = \sum_{i=1}^N \alpha_i \mathbf{e}_i,$$

ahol  $\{\mathbf{e}_i, i = 1, \dots, N\}$  az  $\mathbb{R}^N$  szokásos bázisa. Ekkor

$$\mathbf{p}^{(0)} \mathbf{T}^n = \sum_{i=1}^N \alpha_i \mathbf{e}_i \mathbf{T}^n.$$

Így 4.1. Tétel alapján

$$\lim_{n \rightarrow \infty} \mathbf{p}^{(0)} \mathbf{T}^n = \sum_{i=1}^N \alpha_i \mathbf{e}_i \mathbf{T}^* = \mathbf{p}^* \sum_{i=1}^N \alpha_i,$$

azaz az iteráció ezeket a vektorokat határértékben a  $\mathbf{p}^*$  által kifeszített altérbe viszi át. Látható továbbá az is, hogy azokat a vektorokat, amelyekre  $\sum_{i=1}^N \alpha_i = 0$ ,

az iteráció határértékben a nullvektorba képezi le. Ezért, ha a Markov-láncok definíciójából elhagyjuk a kezdeti vektorra vonatkozó feltételeket, akkor az 4.1. Állítás a következőképp szól: Egy  $(\mathbf{T}, \mathbf{p}^0, S)$  hármas pontosan akkor „eloszlás-ergodikus”, ha  $\mathbf{p}^{(0)}$  megválasztásától függetlenül léteznek a  $\lim_{n \rightarrow \infty} \mathbf{p}^{(n)}$  határértékek és ezek egy adott  $\mathbf{p}^*$  vektor számszorosával egyeznek meg.

**4.2. ÁLLÍTÁS (Markov) [7], [8].** Egy  $(\mathbf{T}, \mathbf{p}^{(0)}, S)$  Markov-lánc pontosan akkor eloszlás-ergodikus, ha  $\mathbf{T}$  mátrixnak létezik egy olyan  $\mathbf{T}^\nu$  hatványa, amelynek legalább egy oszlopában minden elem pozitív.

Ez a feltétel szemléletesen azt jelenti, hogy van olyan  $\nu$  lépésszám, és legalább egy olyan állapot, amelybe  $\nu$  lépéssel minden más állapotból pozitív valószínűséggel eljuthatunk.

A Markov-láncok ergodicitását a  $\mathbf{T}$  mátrix sajátértékeinek segítségével is lehet jellemezni.

**4.1. TÉTEL.** Egy  $(\mathbf{T}, \mathbf{p}^{(0)}, S)$  Markov lánc pontosan akkor eloszlás-ergodikus, ha a  $\mathbf{T}$  mátrixnak a  $\lambda = 1$  az egyetlen egységnyi sajátértéke, és ez egyszeres sajátérték.

*Bizonyítás.* Először tegyük fel, hogy a Markov-lánc eloszlás-ergodikus. Ekkor 4.1. Állítás alapján  $\mathbf{T}$  szemikonvergens mátrix. Mivel  $\mathbf{T}$  sztochasztikus mátrix, ezért 2.8. Állítás alapján a  $\lambda = 1$  az egyetlen egységnyi sajátértéke. Végül 2.7. Állítás és 4.1. Definíció alapján az  $\mathbf{x}\mathbf{T} = \mathbf{x}$  egyenletnek egyetlen megoldása van, ugyanis 4.1. Definíció alapján az 2.7. Állításbeli  $\mathbf{x} = \mathbf{x}(\mathbf{x}^{(0)})$  határvektor független  $\mathbf{x}^{(0)}$ -tól. Ezért a  $\lambda = 1$  egyszeres sajátérték. Most tegyük fel, hogy  $\mathbf{T}$ -nek a  $\lambda = 1$  az egyetlen egységnyi sajátértéke, és ez egyszeres. Ekkor 2.8. Állítás alapján  $\mathbf{T}$  szemikonvergens mátrix. Így 2.7. Állítás és 4.1. Definíció alapján a Markov-lánc eloszlás-ergodikus.

*Megjegyzés.* Vegyük észre, hogy az eloszlás-ergodikus Markov-láncok azonosíthatók a  $H_1 = \{(x_1, x_2, \dots, x_N) \in \mathbb{R}^N \mid x_i \geq 0, \sum_{i=1}^N x_i = 1\}$  halmazon konvergens iterációkkal.

Ezek után jellemezzük a határeloszlást (ha létezik) a  $\mathbf{T}$  átmenetvalószínűségek mátrixának bizonyos tulajdonságai ismeretében. Az első esetben legyen  $\mathbf{T}$  reducibilis mátrix.

1. Tegyük fel, hogy  $\mathbf{T}$ -nek a  $\lambda = 1$  az egyetlen egységnyi sajátértéke és ez egyszeres sajátérték. Ekkor a Markov-lánc eloszlás-ergodikus. Emellett 2.3. Állítás alapján a határeloszlásnak 0 elemei is vannak (azaz  $\mathbf{T}$  határmátrixának van nulla oszlopa).
2. Tegyük fel, hogy  $\mathbf{T}$ -nek a  $\lambda = 1$  az egyetlen egységnyi sajátértéke, de ez nem egyszeres sajátérték. Ekkor a Markov-lánc nem eloszlás-ergodikus. Ugyanakkor 2.8. Állítás és 2.7. Állítás szerint létezik határeloszlása és az függ a kezdeti eloszlástól. Emellett, a határeloszlásnak nulla elemei is vannak.

3. Végül tegyük fel, hogy  $\mathbf{T}$ -nek a  $\lambda = 1$  nem az egyetlen egységnyi sajátértéke. Ekkor a Markov-lánc nem eloszlás-ergodikus és nem létezik határeloszlása sem. (Kivéve, ha a  $\mathbf{T}$  mátrixnak az egyhez tartozó baloldali sajátvektorából indítjuk az iterációt).

Legyen most  $\mathbf{T}$  irreducibilis.

1. Tegyük fel, hogy a  $\mathbf{T}$  mátrixnak a  $\lambda = 1$  az egyetlen egységnyi sajátértéke és ez egyszeres sajátérték. Ekkor a Markov-lánc eloszlás-ergodikus. Emellett 2.1. Állítás és 2.7. Állítás miatt a határeloszlás minden eleme pozitív. Ennek az 2.4. Állítás alapján a  $\text{tr } \mathbf{T} > 0$  egy elégséges feltétele.
2. Ha  $\mathbf{T}$  irreducibilis, akkor 2.1. Állítás miatt a  $\lambda = 1$  mindenképpen egyszeres sajátérték, ezért ekkor a reducibilis mátrixok esetében felsorolt második eset nem fordul elő.
3. Végül tegyük fel, hogy a  $\mathbf{T}$  mátrixnak a  $\lambda = 1$  nem az egyetlen egységnyi sajátértéke. (Ennek a 2.4. Állítás alapján a  $\text{tr } \mathbf{T} = 0$  egy szükséges feltétele). Ekkor a Markov-lánc nem eloszlás-ergodikus és nem létezik határeloszlása sem. (Kivéve, ha a  $\mathbf{T}$  mátrixnak az egyhez tartozó baloldali sajátvektorából indítjuk az iterációt.)

## 5. Osztály-ergodikus Markov-láncok

Legyen  $(\mathbf{T}, \mathbf{p}^{(0)}, S)$  egy Markov-lánc.

**5.1. Definíció.** Azt mondjuk, hogy az  $s_j$  állapot elérhető  $s_i$ -ből, ha véges számú lépés alatt eljuthatunk  $s_i$ -ből  $s_j$ -be.

**Jelölés:**  $s_i \rightarrow s_j$ .

**5.2. Definíció.** Azt mondjuk, hogy  $s_i$  és  $s_j$  érintkeznek, ha  $s_i \rightarrow s_j$  és  $s_j \rightarrow s_i$ .

**Jelölés:**  $s_i \leftrightarrow s_j$

Az érintkezés az állapottéren értelmezett ekvivalencia reláció, ezért azt ekvivalenciaosztályokra bontja fel.

**5.3. Definíció.** A Markov-lánc osztályainak az érintkezési reláció által indukált ekvivalenciaosztályokat nevezzük. Azt mondjuk, hogy egy  $\beta$  osztály elérhető az  $\alpha$  osztályból, ha léteznek olyan  $s_i \in \alpha$  és  $s_j \in \beta$  állapotok, amelyekre  $s_i \rightarrow s_j$ . Egy osztályt végesnek nevezünk, ha más osztály nem érhető el belőle. Ha egy véges osztály egyetlen állapotot tartalmaz, akkor azt az állapotot elnyelőnek hívjuk.

**5.4. Definíció.** Az  $s_i$  állapotot átmenetinek nevezzük, ha létezik olyan  $s_j$ , hogy  $s_i \rightarrow s_j$ , de  $s_j \nrightarrow s_i$ . Egyébként  $s_i$ -t ergodikusnak nevezzük. (Azaz,  $s_i$  pontosan akkor ergodikus, ha  $s_i \rightarrow s_j$  esetén  $s_j \rightarrow s_i$ .)



Ezért, ha a Markov-lánc egy osztályában egy állapot átmeneti (ergodikus), akkor az osztályban lévő tetszőleges állapot átmeneti (ergodikus). Ez motiválja a következő definíciót:

**5.5. Definíció.** A Markov-lánc egy  $\alpha$  osztályát átmenetinek nevezzük, ha tartalmaz átmeneti állapotot, egyébként ergodikusként nevezzük.

Nyilvánvalóan egy osztály pontosan akkor ergodikus, ha véges.

**5.6. Definíció.** A Markov-láncot osztály-ergodikusként nevezzük, ha az állapot-teret egyetlen ergodikus osztályból áll. Regulárisnak nevezzük, ha egy rögzített  $k$  lépésszámnál pozitív valószínűséggel juthatunk el bármelyik állapotból bármelyik állapotba. Periodikusnak nevezzük, ha osztály-ergodikus, de nem reguláris.

**5.1. ÁLLÍTÁS.** Legyen  $T$  egy Markov-lánc átmenetvalószínűségek mátrixa. Ekkor a lánc:

1. pontosan akkor osztály-ergodikus, ha  $T$  irreducibilis,
2. pontosan akkor reguláris, ha  $T$  primitív,
3. pontosan akkor periodikus, ha  $T$  imprimitív.

*Bizonyítás.*

1. A 2.5. Állítás alapján  $T$  pontosan akkor irreducibilis, amikor  $G(T)$  erősen összefüggő. Ez pontosan akkor igaz, amikor a Markov-lánc állapotai elérhetők egymásból, azaz pontosan akkor, amikor a Markov-lánc osztály-ergodikus.
2. A 2.3. Állításból azonnal következik, mivel egy Markov-lánc pontosan akkor reguláris, amikor az átmenetvalószínűségek mátrixának van pozitív hatványa (azaz ha egy rögzített  $k$  lépésszámnál pozitív valószínűséggel juthatunk el bármelyik állapotból bármelyik állapotba).
3. Közvetlenül következik 2.3. Definícióból és 5.6. Definícióból.

*Megjegyzés.* Vegyük észre, hogy azok a Markov-láncok, amelyek eloszlás- és osztály-ergodikusak is egyben, bármely valószínűségeloszlásból indítva, mindig a  $\lambda = 1$  sajátértékhez tartozó valószínűségeloszlásba (sajátvektorba) jutnak. (Ez igaz valamennyi eloszlás-ergodikus Markov-láncre is). Ezen Markov-láncok stacionárius eloszlását ezért stabilnak is nevezhetnénk. Azon Markov-láncokat, amelyek osztály-ergodikusak, de nem eloszlás-ergodikusak, ha nem a  $\lambda = 1$  sajátértékhez tartozó valószínűségeloszlásból (sajátvektorból) indítjuk, akkor a valószínűségeloszlásnak nincs határértéke, így a stacionárius eloszlást instabilnak is nevezhetnénk. Ez lehetőséget adna egy újabb ergodicitási fogalom bevezetésére is, amely a stacionárius eloszlás stabilitásán alapulhatna.

## 6. Példák ergodikus Markov-láncokra

**6.1. Példa.** Tekintsünk egy olyan három állapotú rendszert, amelyben cirkuláris átmenet lehetséges, s minden lépés egyirányú. Ez egy olyan atom modelljének tekinthető, amely az alapállapotból külső hatásra az egyik gerjesztett állapotba juthat, ahonnan egy közbülső energiaszint érintésével bomlik le (lézer). A rendszer átmenetvalószínűségek mátrixa:

$$\mathbf{T} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

A  $\mathbf{T}$  mátrix erősen összefüggő, ezért irreducibilis. Így 5.1. Állítás szerint a Markov-lánc osztály-ergodikus. A  $\mathbf{T}$ -hez rendelt irányított gráf alapján (5. ábra)  $\mathbf{T}$  egy három-indexű imprimitív mátrix, ezért a Markov-lánc nem eloszlás-ergodikus.

**6.2. Példa.** Tekintsünk egy diszkrét egydimenziós diffúziós folyamatot, periodikus peremfeltételekkel. A rendszer álljon most  $N$  cellából, és legyen a diffúziós együttható  $D = \text{konstans} < 0,5$ . A rendszer átmenetvalószínűségek mátrixa most azt mutatja meg, hogy egységnyi idő alatt az adott cella tartalmának mekkora része kerül a szomszédos cellákba. A Markov-lánc átmenetvalószínűségek mátrixa:

$$\mathbf{T} = \begin{pmatrix} 1-2D & D & 0 & \cdot & \cdot & \cdot & D \\ D & 1-2D & D & \cdot & \cdot & \cdot & 0 \\ 0 & D & 1-2D & D & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & D & 1-2D & D \\ D & \cdot & \cdot & \cdot & 0 & D & 1-2D \end{pmatrix}.$$

A  $\mathbf{T}$  mátrixhoz rendelt irányított gráf erősen összefüggő (6. ábra), ezért a mátrix irreducibilis, és mivel  $\text{tr } \mathbf{T} > 0$ , ezért a mátrix primitív. Tehát a Markov-lánc egyaránt osztály- és eloszlás-ergodikus is.

**6.3. Példa.** Tekintsük megint az előző rendszert, és most azt tesszük fel, hogy bizonyos helyeken advekcio (drift) is van, méghozzá  $v = 2D$  illetve  $v = -2D$  nagyságú. Ekkor a Markov-lánc átmenetvalószínűségek mátrixa:

$$\mathbf{T} = \begin{pmatrix} 1-2D & D & 0 & \cdot & \cdot & \cdot & D \\ D & 1-2D & D & \cdot & \cdot & \cdot & 0 \\ 0 & D+0,5|v| & 1-2D & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & 1-2D & D+0,5|v| & \cdot \\ 0 & \cdot & \cdot & \cdot & D & 1-2D & D \\ D & \cdot & \cdot & \cdot & 0 & D & 1-2D \end{pmatrix}.$$

A  $T$  mátrixhoz rendelt irányított gráf nem erősen összefüggő (7. ábra), ezért a mátrix reducibilis, így a Markov-lánc nem osztály-ergodikus. Ha a  $T$  mátrixot reducibilis alakra hozzuk, nyilvánvaló, hogy a Markov-lánc eloszlás-ergodikus.

### Irodalom

- [1] A. Berman and R. J. Plemmons, *Nonnegative matrices in the mathematical sciences*, Academic Press (New York, 1979).
- [2] I. Faragó and M. Kovács, Stochastic regular splitting and its application to the iterative methods, *Pure Math. Appl.* (megjelenés alatt).
- [3] F. R. Gantmacher, *Matrizenrechnung*, Deutscher Verlag der Wissenschaften (Berlin, 1966).
- [4] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge Univ. Press (London, 1986).
- [5] S. Karlin and H. M. Taylor, *Sztochasztikus folyamatok*, Gondolat (Budapest, 1985).
- [6] B. Philippe and Y. Saad, Numerical method in Markov-chain modeling, *Operations Research*, 40 (1992), 1156–1179.
- [7] A. Prékopa, *Valószínűségelmélet*, Műszaki könyvkiadó (Budapest, 1980).
- [8] A. Rényi, *Valószínűségszámítás*, Tankönyvkiadó (Budapest, 1954).
- [9] F. S. Roberts, *Discrete mathematical models, with applications to social, biological and environmental problems*, Prentice Hall (Englewood Cliffs, 1976).
- [10] P. Rózsa, *Lineáris algebra és alkalmazásai*, Műszaki Könyvkiadó (Budapest, 1976).
- [11] P. Semal, Iterative algorithms for large stochastic matrices, *Linear algebra and its application*, 154 (1991), 65–103.
- [12] R. Varga, *Matrix Iterative Analysis*, Prentice Hall (Englewood Cliffs, 1962).

(Beérkezett: 2000. május 18.)

FARAGÓ ISTVÁN  
EÖTVÖS LORÁND TUDOMÁNYEGYETEM  
ALKALMAZOTT ANALÍZIS TANSZÉK  
BUDAPEST, KECSKEMÉTI U. 10  
H-1053  
E-MAIL: FARAGOIS@CS.ELTE.HU

KOVÁCS MIHÁLY  
EÖTVÖS LORÁND TUDOMÁNYEGYETEM  
ALKALMAZOTT ANALÍZIS TANSZÉK  
BUDAPEST, KECSKEMÉTI U. 10  
H-1053  
E-MAIL: KMISI@CS.ELTE.HU

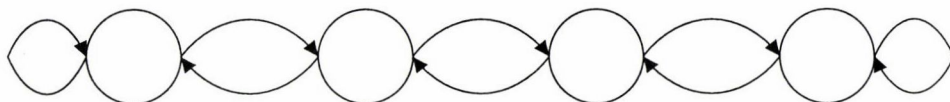
## ERGODIC MARKOV CHAINS AND THEIR APPLICATIONS

ISTVÁN FARAGÓ AND MIHÁLY KOVÁCS

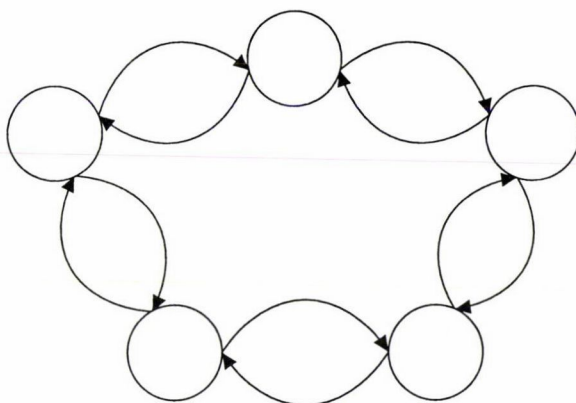
Both the Markov processes and the Markov chains play a fundamental role in the different topics of the natural sciences. Considering the Markov chains, one of the basic problems in the characterization of a system after a long time period is whether it is possible to observe any "stabilization" around a well-defined state. In this paper we analyse the scalar ergodic Markov chains. Our investigation is based on the one-step iterative methods of numerical linear algebra. Moreover, we give some important physical applications of the theory, too.



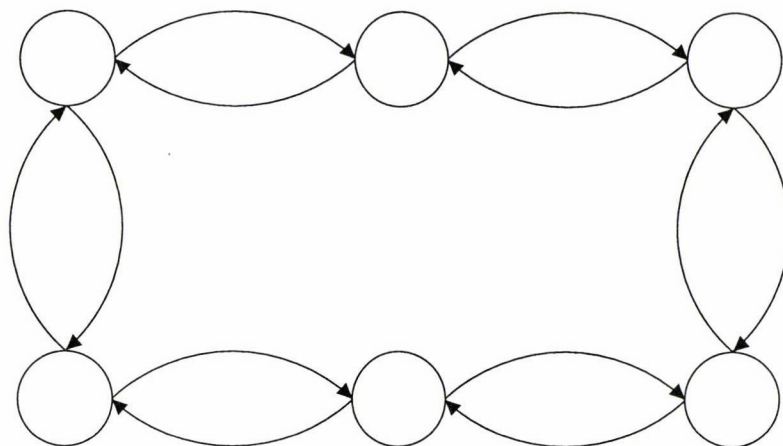
1. ábra



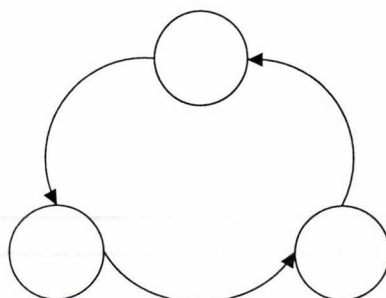
2. ábra



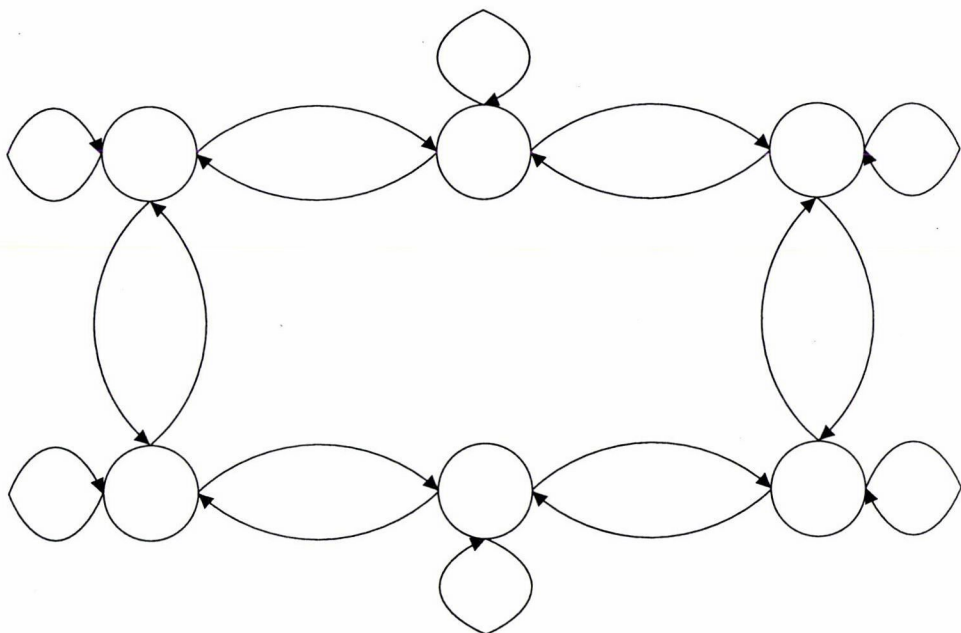
3. ábra



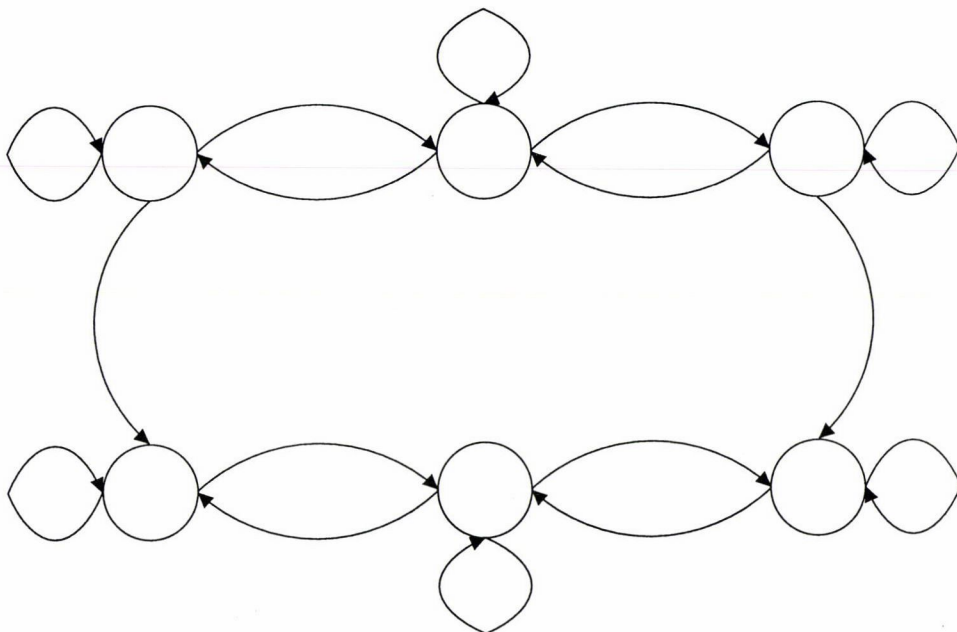
4. ábra



5. ábra



6. ábra



7. ábra

## VÁZKIJELÖLÉS KONTÚR-REPREZENTÁCIÓJÚ KÉPEKEN

FAZEKAS ATTILA ÉS SÁNTA ISTVÁN

Debrecen

A digitális képfeldolgozásban vonalas ábrák feldolgozása során gyakran használatos technika a vázkijelölés. A legtöbb, az irodalomból ismert algoritmus ún. pixel-pixel reprezentációjú képeken dolgozik. Ebben a cikkben egy olyan vázkijelölő algoritmust konstruálunk Arcelli [1] algoritmusának alapelveit felhasználva, amely kontúr-reprezentációjú képeken dolgozik.

### Bevezetés

A digitális képfeldolgozásban vonalas ábrák feldolgozása során gyakran használatos technika a vázkijelölés. A vázkijelölés pontos fogalmának ismerete nélkül ezen eljárások szemléletes hatása az, hogy egy objektumot egy olyan egy pixel vastag „középvonallal” helyettesítenek, amelynek kapcsolódási viszonyai megegyeznek az eredeti objektum kapcsolódási viszonyaival. A váz pontos, matematikai fogalmát és legfontosabb tulajdonságait karakterizáló elméleti eredményeket [7] alapján, az alapvető vázkijelölési technikák összefoglaló elemzését [4] értelmezésében használjuk.

Az irodalomban a digitális képek reprezentálására számos módszer ismeretes. A leggyakoribb az ún. pixel-pixel reprezentáció, amikor a képet alkotó képpontok geometriai elhelyezkedését és világosságkódját egy mátrixszal reprezentáljuk. Ezt a tényt elsősorban az magyarázza, hogy a rendelkezésre álló hardver elemek szinte mindegyike ezt a reprezentációt támogatja. A vázkijelölő algoritmusokra is – mint speciális képfeldolgozásbeli eljárásokra – igaz a fenti megállapítás. Bár az irodalomból ismeretes olyan algoritmus, amely futam-hossz reprezentációjú képen [5], illetve vetületi képeken [3] végzi el a vázkijelölést. Az elemek mátrixbeli koordinátája a képpontok geometriai elhelyezkedését, értéke pedig a képpontok világosságkódját jelöli. A pixel-pixel reprezentációkon kívül még használatos a futam-hossz [5], a 4-fa [2] és a kontúr-reprezentáció [1] is. Az orvosi képfeldolgozásban gyakran a bináris képek merőleges vetületei állnak csak rendelkezésre. Bizonyos feltételek teljesülése esetén ezek a vetületek is az adott kép reprezentációjának tekinthetők [3].

Bináris képek esetén a kontúr-reprezentáció alkalmazásával általában jelentős mértékben csökkenthetjük a kép tárolásához szükséges memória méretét, hiszen az objektumok tárolása helyett csak azok kontúrjait és a kontúrok egymáshoz való viszonyát kell tárolnunk. Dolgozatunkban egy olyan algoritmus elkészítését tűztük ki célul, amely kontúr-reprezentációjú bináris képen üzemel. Az algoritmusunk Arcelli egyik algoritmusának [1] kontúr-reprezentációra történő adaptálásával készült.

## 1. Bináris képek kontúr-reprezentációja

A kontúr-reprezentáció alapötlete az, hogy az eredeti bináris kép tárolása helyett elegendő a képen előforduló kontúrokat és azok egymáshoz való viszonyát tárolni. Ismeretes, hogy egy nemüres zárt objektum kontúrja egy nemüres zárt görbe lesz [1]. Ez azt jelenti, hogy egy kontúrt tárolhatunk olyan módon, hogy a kontúrt alkotó zárt görbe egy tetszőleges pontjának (kezdőpont) abszolút koordinátáját és a többi kontúrponthoz viszonyított – relatív koordinátáját tároljuk. A relatív koordináták legalkalmasabb kódolási módszere az ún. Freeman-féle lánckód [6].

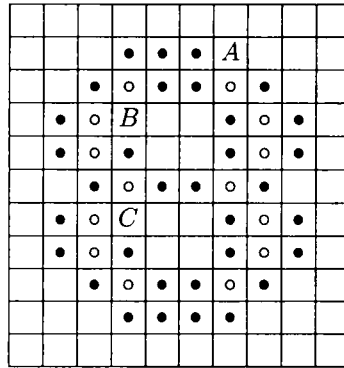
Könnyen belátható, hogy egy kontúr körbezárhat, vagy részben átfedhet, de nem metszhet egyetlen más kontúrt sem. A kontúrok ezen tulajdonsága alapján építünk fel egy tartalmazási fát. A fa gyökere a teljes bináris képet, amíg a többi csúcs egy-egy kontúrt reprezentál. Ha a gyöktől valamely levélelemig haladó úton egy csúcs közelebb van a gyökérelemhez, mint egy másik, akkor a közelebbi csúcs által reprezentált kontúr tartalmazza – körbezárja – a másik csúcs által reprezentált kontúrt. Az 1. ábrán egy objektum, amíg a 2. ábrán annak tartalmazási fája látható. Az 1. ábrán a nem jelölt pontok a háttérpontokat, a  $\bullet$ -tal jelöltek a kontúr-pontokat, amíg a  $\circ$ -rel jelöltek az objektum belső pontjait jelölik. Az ábra példát mutat arra az esetre is, amikor két kontúr (a  $B$  és a  $C$ ) részben átfedi egymást. Betűkkel jelöltük a kontúrok kezdőpontját. A 2. ábrán ezen betűkkel hivatkozunk az egyes kontúrokra.

A gyöktől számítva páratlan szinten álló csúcsok által reprezentált kontúrokat külső, amíg a páros szinten álló csúcsok által reprezentált kontúrokat belső kontúroknak nevezzük. Megállapodás alapján a külső kontúrokat az óramutató járásával ellentétes irányban, amíg a belső kontúrokat az óramutató járásával megegyező irányban járjuk be.

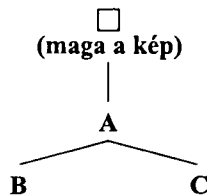
Ha ismerjük egy bináris kép kontúr-reprezentációját, akkor a kép pixel-pixel reprezentációja könnyen meghatározható. Ugyanis a kontúrok tartalmazási relációját reprezentáló fa gyökéreleméből kiinduló, valamely levélelemig haladó utakat alkotó csúcsokhoz tartozó kontúrokat a gyökérelemtől távolodó sorrendben meg-rajzoljuk, majd az általuk körbehatárolt képterületet külső kontúrok esetén előtér színre, belső kontúrok esetén háttér színre színezzük.

Ha a pixel-pixel és a kontúr-reprezentáció közötti konverziók során a kép eredeti fizikai méretét is meg szeretnénk őrizni, akkor azt külön kell tárolnunk. A kontúr-





1. ábra. Egy bináris objektum és kontúrjai



2. ábra. Az 1. ábrán látható objektum kontúrjainak tartalmazási fája

reprezentáció többszintű képek esetén is alkalmazható azzal a módosítással, hogy az egyes kontúrokat reprezentáló csúcsokhoz az általuk körbehatárolt képterület világosságkódját is hozzá kell rendelni.

## 2. A kontúr-reprezentáció tárolására szolgáló adatszerkezet

A kontúr-reprezentáció tárolására az adatszerkezetet úgy választottuk meg, hogy az egyaránt lehetővé tegye a kontúrok memóriában és állományban történő tárolását is.

Állományban a képet meghatározó kontúrok mellett az azok egymáshoz való viszonyát reprezentáló tartalmazási fa, továbbá az előző fejezetben említett okból, a kép fizikai mérete is tárolásra kerül. Az állomány kényelmes kezelése érdekében lehetővé tesszük egy hossz-adat bevezetésével azt, hogy a tartalmazási fát tartalmazó állományrészt figyelmen kívül hagyhassuk. A tartalmazási fa ismerete a konverzió megvalósításához elengedhetetlen. A fát a klasszikus preorder bejárás szerint tároljuk, de tetszőleges ezzel egyenértékű más tárolási mód is alkalmazható.

Minden kontúrt listaként tárolunk. A kontúr kezdőpontjának abszolút koordinátáit a lista fejében tároljuk a következő formában:

fejall=koord[1..2] of word;

majd következnek a kontúrponthoz tartozó listaelemek (fejjel kezdődően):

```
lmall=record
    elozo,kov:byte;
end;
```

A két mező értelemszerűen a bejárás szempontjából a megelőző és a rákövetkező kontúrponthoz tartozó relatív koordinátáját adja meg – Freeman-féle lánckód formájában – a vizsgált kontúrponthoz viszonyítva (3. ábra).

3	2	1
4	P	0
5	6	7

3. ábra. A *P* ponthoz tartozó Freeman-féle lánckód a 8-szomszédjainak függvényében

A fenti adatszerkezeteket felhasználva a memóriában a kontúrt egy *lmtip* típusú listaként tároljuk. Ehhez egy *kovlm* mutatómezőt vezetünk be (ez egy *lmtip* típusú rekordra mutat – erre utal a típusnév is: *lmmut*). Az algoritmus hatékony implementálása érdekében alkalmaztunk még két további mezőt. Az egyik azt mondja meg, hogy hányszor jutottunk el az adott pontba a kontúrkövetés során, ezt *számosságnak* neveztük (*szamos*). A másik egy logikai mező, amely a vázponthoz megkülönböztetésére szolgál majd (*vaze*).

```
lmtip=record
    adat:lmall;
    vaze:boolean;
    szamos:byte;
    kovlm:lmmut;
end;
```

A fejek koordinátáit és a hozzájuk tartozó listaelem mutatót egy *fejtip* típusú tömbbe helyezzük el:

```
fejtip=record
    adat:fejall;
    kezd:lmmut;
end;
```

A *kezd* a kontúrt reprezentáló lista kezdőcíme.

Külön említést érdemel az a speciális eset, amikor a kontúr egy pixel hosszú. Ezt a memóriában egy olyan egyelemű listaként reprezentáljuk, aminek első és egyetlen eleme önmagára mutat (a *kovlm* mezőjével). Így ha a feldolgozás során egy hosszabb kontúrból egy egy pixel hosszú kontúr keletkezik, akkor az algoritmus egy fent leírt tulajdonságú listát fog létrehozni.

### 3. Vázkijelölő algoritmus

Ebben a fejezetben az általunk kifejlesztett vázkijelölő algoritmust ismertetjük, amely kontúr-reprezentációjú képeken üzemel. Algoritmusunk – ahogy korábban már jeleztük – Arcelli [1] kontúrszekvenciális algoritmusán alapszik.

A kontúr pontjainak abszolút koordinátáit nem tároljuk, mivel a kezdőpont abszolút koordinátájából és a kontúrt reprezentáló listából könnyen meghatározhatók. Hatékony végrehajtás érdekében felépítettünk egy táblázatot, amiben a kontúrpontok abszolút koordinátái szerepelnek olyan sorrendben, ahogy azok a kontúrban elhelyezkednek.

Arcelli alapján algoritmusunk iteratív lesz. Minden iteratív menet két almenetre bomlik. Az elsőben pontonként megvizsgáljuk a kontúrokat és megjelöljük azokat a kontúrpontokat, amelyek vázpontok lesznek. A másodikban minden olyan kontúrpontot törölünk, ami nem vázpont. A törölt kontúrpontok helyett új belső pontokat veszünk fel a kontúrba.

Az iterációs menetek mindaddig ismétlődnek, amíg van törlendő pont a kontúrban. Az iteráció végén egy speciális kontúrt kapunk, ami tulajdonképpen maga a váz lesz.

#### 3.1. A vázpontok kijelölése

Egy  $P$  kontúrpont  $C$ -szomszédain azokat a kontúrpontokat értjük, amelyek  $P$  – a kontúrkövetés sorrendjében – közvetlen szomszédjai a kontúrban. Azaz egy tetszőleges kontúrpont valamelyik kontúrbeli adott előfordulásának pontosan két  $C$ -szomszédja van, az egyik megelőzi, a másik követi az adott előfordulást az adott kontúrban. A  $C$ -szomszéd fogalom fontos szerepet fog játszani a későbbiekben, mert egy pont előfordulhat többször is egy adott kontúrban, illetve több kontúrban is szerepelhet.

Azokat a kontúrpontokat minősítjük vázponttá, amelyek többszörös pontok. A többszörös pont fogalmát Arcelli alapján a következőképpen adjuk meg:

*3.1. Definíció* [1]. Egy objektumpontot többszörösnek nevezünk, ha a következő feltételek közül legalább egyet teljesít:

- (1) a kontúrkövetés során többször áthaladtunk rajta,
- (2) nincs szomszédja az objektum belsejében,
- (3) van legalább egy közvetlen szomszédja, amelyik kontúrpont, de nem  $C$ -szomszédja.

**3.2. JELÖLÉS.** Legyen  $P, Q$  két 8-szomszédos pont. Jelölje  $\gamma_P(Q)$  azt az  $i \in [0, 7]$  egész számot, amely azt mutatja, hogy  $Q$  milyen irányban helyezkedik el  $P$ -hez képest a Freeman-féle lánckód szerint. Ekkor a  $\gamma_P^{-1}(i)$ -vel jelölt inverzfüggvény értéke maga a  $Q$  pont lesz.

A 3.1. Definíció (1)-feltételének vizsgálata miatt alkalmaztuk a *számosság* nevű mezőt az adatszerkezetben. Ennek meghatározásához végig kell mennünk a kontúrokon.

A 3.1. Definíció (2)-feltételének vizsgálatához a vizsgált kontúrponthoz (jelöljük  $P$ -vel)  $C$ -szomszédjainak  $P$ -hez, illetve egymáshoz viszonyított helyzetét használjuk fel, ami meghatározza, hogy az adott kontúrponthoz van-e szomszédja az objektum belsejében. Tegyük fel, hogy  $P$  megelőző  $C$ -szomszédja  $E$ , rákövetkezője  $K$ , illetve  $\gamma_P(K) = i$  és  $\gamma_P(E) = j$ . Ebben az esetben számunkra a  $\gamma_P^{-1}((i+1) \bmod 8)$  és  $\gamma_P^{-1}((j-1) \bmod 8)$  közötti pontok fontosak, amiket mi *belső oldali* szomszédoknak neveztünk el, hiszen ha van a vizsgált pontnak olyan szomszédja, amely belső pont, az csak ezek közül kerülhet ki. Ennek illusztrálására szolgál a 4. ábra, amelyen a betűk a fent említett pontokat jelentik. Ekkor a  $\circ$ -rel jelölt pontok világossághódja számunkra nem fontos, mivel nem tartoznak bele  $P$  8-szomszédjába. A nem jelölt pontok biztos, hogy háttérpontok lesznek, tehát a feltételek teljesülésének ellenőrzésekor szintén nem kell vizsgálnunk őket. A  $\bullet$ -tal jelölt pontok viszont biztosan objektumpontok lesznek (kivéve ha egy pixel széles az objektum ezen a részen), így ezek fogják meghatározni, hogy teljesül-e a (2)-feltétel. Látható, hogy ha  $\gamma_P(K) = i$ , illetve  $\gamma_P(E) = j$ , akkor bármely  $\bullet$ -tal jelölt pont esetén – amelyhez a  $\gamma$  függvény a  $k$ -t rendeli értéként – teljesül az  $i+1 \leq k \leq j-1$  egyenlőtlenség.

$\circ$				$\circ$
$\circ$	$K$	$P$	$E$	$\circ$
$\circ$	$\bullet$	$\bullet$	$\bullet$	$\circ$
$\circ$	$\circ$	$\circ$	$\circ$	$\circ$

4. ábra. A (2)-feltétel teljesüléséhez csak a  $\bullet$ -tal jelölt pontokat kell ellenőrizni

Könnyű igazolni, hogy a fenti állítás a  $P$ ,  $E$ ,  $K$  pontok bármely konfigurációjára igaz.

Meg kell még vizsgálnunk a lehetséges konfigurációkat abból a szempontból, hogyan fordulhat elő, hogy a belső oldali szomszédok közül egy sem belső pont.

A legegyszerűbb eset, amikor nincs belső oldali szomszéd. Ez akkor igaz, ha  $\gamma_P(K) = \gamma_P(E)$ , vagy  $\gamma_P(K) = (\gamma_P(E) - 1) \bmod 8$ . Ekkor teljesül az 3.1. Definíció (2)-feltétele. Másik eset, amikor a  $\gamma_P(K) = (\gamma_P(E) - 2) \bmod 8$ , amint az 5. ábrán látható, amelyen a 4. ábra jelöléseit használtuk.

$\circ$				$\circ$
$\circ$		$P$	$E$	$\circ$
$\circ$	$\circ$	$K$	$\bullet$	$\circ$
$\circ$	$\circ$	$\circ$	$\circ$	$\circ$

5. ábra. A  $\gamma_P(K) = \gamma_P(E) - 2$  esethez tartozó konfiguráció

Amikor tehát a pontok az 5. ábrának vagy annak  $k \cdot 90^\circ$ -os ( $k \in \mathbb{Z}$ ) elforgatottjának megfelelő konfigurációban helyezkednek el, akkor két lehetősége van annak, hogy a  $P$ -nek ne legyen belső szomszédja. Az egyik esetben a  $\bullet$ -tal jelölt pont szerepel egy kontúrban, ekkor ugyanis nem lesz a  $P$  körül egyetlen olyan pont sem,

ami belső pont lehetne. A másik esethez meg kell jegyeznünk, hogy a fent említett konfiguráció csak külső kontúrban fordulhat elő. Lehetséges ugyanis, hogy  $K$  úgy szerepel egy kontúrban, hogy azon előfordulásának rákövetkezője  $E$  lesz. Ez pedig csak akkor teljesül, ha az utóbb említett kontúr egy belső kontúr, vagy a  $P$ -t tartalmazó kontúrnak a másik oldala. De akár egy másik belső kontúrról, akár ugyanarról a külső kontúrról van is szó, az biztos, hogy a  $\bullet$ -tal jelölt pont háttérpont lesz, tehát  $P$ -nek nincs olyan szomszédja, ami belső pont lehetne.

A többi lehetséges konfigurációt tekintve  $((\gamma_P(E) - \gamma_P(K)) \bmod 8 > 2)$  könnyen látható, hogy ahhoz, hogy a (2)-feltétel teljesüljön, teljesülnie kell a (3)-feltételnek is, így elegendő azt vizsgálni.

Az 3.1. Definíció (3)-feltételének ellenőrzéséhez szintén meg kell vizsgálnunk a kontúrokat alkotó objektumpontokat. Most olyan belső oldali 4-szomszédokat kell keresnünk, amelyek szerepelnek valamelyik kontúrban. Ha egy-egy ilyen pont előfordulásainak  $C$ -szomszédjai között sehol sem szerepel  $P$ , akkor  $P$  kielégíti a (3)-feltételt.

A feltételek ellenőrzéséhez elengedhetetlen az abszolút koordináták ismerete, amit könnyen kiszámíthatunk a kontúr kezdőpontjának és az adott kontúrt reprezentáló listának az ismeretében. Ezt a végrehajtás hatékonysága szempontjából célszerű az algoritmus inicializáló lépésében megtenni és táblázatban tárolni.

### 3.2. A nem-vázpontok törlése

Vázkijelölő algoritmusunk a törlendő kontúrponthoz a kontúrkövetés sorrendjében egyenként törli. A törlés egyik alapvető nehézségét az jelenti, hogy karakterizálnunk kell azokat az objektumpontokat, amelyeket a törölt kontúrponthoz helyére a kontúrba fel kell vennünk. A következőkben az ezen kérdés megválaszolását biztosító elméleti eredményeinket fogalmazzuk meg. Az első két tétel a törlés megvalósításának módját adja meg.

**3.3. TÉTEL.** *Legyen  $O$  egy objektum,  $L$  annak 8-kapcsolódó kontúrja. Az  $O$  objektum  $L$  kontúrját alkotó objektumpontok valamelyikének törlésével előáll  $O'$  objektum  $L'$  kontúrját úgy határozhatjuk meg, hogy a törölt objektumpontot annak belső oldali 4-szomszédjaival helyettesítjük az  $L'$ -ben.*

*Bizonyítás.* Az  $L$  kontúrt alkotó kontúrponthoz van olyan 4-szomszédjuk, ami háttérpont. A kontúr tetszőleges  $P$  kontúrponthoz pontosan akkor törölhető, ha minden belső oldali 4-szomszédja (ha van) belső pont. Ha ugyanis ezek közül valamelyik háttérpont lenne, akkor többször át kellene haladnunk a  $P$ -n a kontúrkövetés során. Ez pedig a 3.1. Definíció (1)-feltételének értelmében azt jelenti, hogy a  $P$  nem lenne törölhető pont. Ha esetleg valamelyik belső oldali 4-szomszéd kontúrponthoz lenne, akkor a 3.1. Definíció (3)-feltételének értelmében szintén nem lehetne a  $P$  pontot törölni. Felléphet még az az eset is, hogy a  $P$  pontnak nincs is belső oldali objektumpont szomszédja, de akkor a 3.1. Definíció (2)-feltétele nem engedi törölni a  $P$ -t.

Vizsgáljuk meg, hogy milyen hatással van a  $P$  kontúrponthoz törlése a  $P$  8-szomszédosságában található pontokra. Nyilvánvalóan a  $P$  8-szomszédosságában ta-

lálható háttérpontokra nincs hatással a  $P$  pont törlése. Hasonlóképpen a  $P$  pont 8-szomszédságában található kontúrpontra sincs hatással, hiszen azoknak a  $P$  törlése előtt már volt háttérpont a 4-szomszédságukban, ami a törlés után is megmarad. A  $P$  8-szomszédságában található tetszőleges  $Q$  kontúrpontra nem válhat a  $P$  törlése után izolált ponttá. Ez csak akkor következhetne be, ha  $Q$  8-szomszédságában egyetlen objektumpont a  $P$  lenne. Ha  $P$  8-szomszédságában is legfeljebb két objektumpont van, akkor  $P$ -t nem lehetne törölni a 3.1. Definíció (2)-feltétele alapján. Ezek alapján azt is megállapíthatjuk, hogy a  $P$  pont  $C$ -szomszédjai is kontúrpontra maradnak a  $P$  törlése után. A  $P$  4-szomszédságában található belső pontok azonban kontúrpontra válnak, hiszen lesz a 4-szomszédságukban egy háttérpont, a törölt  $P$  pont. Levonhatjuk tehát azt a következtetést, hogy  $P$  törlése csak a belső oldali 4-szomszédjaira van hatással, mégpedig kontúrpontra teszi őket. Ez azt jelenti, hogy ezeket a pontokat kell felvennünk a kontúrba a  $P$  törlése után, annak helyére.  $\square$

**3.4. TÉTEL.** Legyen  $L$  egy 8-kapcsolódó kontúr. A 3.1. Definíció által karakterizált tetszőleges  $P$  pont törlése után a kontúrba nulla vagy legfeljebb három pontot kell felvenni.

*Bizonyítás.* Mivel a törölt  $P$  helyére csak a 4-szomszédjai közül vehetünk fel objektumpontokat, így négynél több objektumpont felvétele lehetetlen. Megmutatjuk, hogy négy objektumpontot sem lehet felvenni. Tegyük fel, hogy négy objektumpontot kell felvennünk. Ez azt jelenti, hogy a  $P$  pont minden 4-szomszédját fel kell venni, de ebben az esetben a  $P$  pontnak nincs olyan 4-szomszédja, ami háttérpont lenne, azaz  $P$  nem kontúrpontra, így nem törölhetjük. Ez pedig ellentmondás.

Az elkövetkezőkben belátjuk, hogy a tétel állításában megfogalmazott esetek előfordulhatnak. Ennek érdekében példát adunk az egyes esetekre, amit ábrákkal is illusztrálunk. Az ábrákon a  $P$  pont a korábbiakkal összhangban a törlendő objektumpontot,  $E$  és  $K$  pedig a megelőző, illetve rákövetkező  $C$ -szomszédjait jelöli. A számokkal a felvételre kerülő objektumpontokat jelöltük, a számozás sorrendje az újonnan előálló kontúrbeli sorrendjüket határozza meg.

A 6. ábra mutatja, hogy van olyan eset, amikor három objektumpontot kell felvennünk a törlendő kontúrpontra helyére.

	$K$		$E$	
	3	$P$	1	
		2		

6. ábra. A  $P$  pont törlése után három új objektumpont kerül a kontúrba

A 7. ábra pedig arra szolgáltat bizonyítékot, hogy létezik olyan eset, amikor két objektumpontot kell felvennünk a törlendő kontúrpontra helyére.

Végül egy objektumpont felvételére látható példa a 8. ábrán.

	<i>K</i>			
	2	<i>P</i>	<i>E</i>	
		1		

7. ábra. A *P* pont törlése után két új objektumpont kerül a kontúrba

	<i>K</i>	<i>P</i>	<i>E</i>	
		1		

8. ábra. A *P* pont törlése után egy új objektumpont kerül a kontúrba

	<i>K</i>	<i>P</i>		
		<i>E</i>		

9. ábra. A *P* pont törlése után nem kerül új objektumpont a kontúrba

Lehetséges olyan eset is, amikor nem kell felvennünk új objektumpontot a kontúrba. Ekkor *E*-nek a rákövetkezője lesz *K*, illetve *K*-nak a megelőzője lesz *E* a *P* törlése után. Erre mutat példát a 9. ábra.

A tétel az összes lehetséges eset leszámolásával is bizonyítható. Ez a leszámolás egy olyan függvény implementálásával valósítható meg, amely meghatározza, hogy az aktuális kontúrpon (amire az aktmut mutat) törlése esetén hány pontot kell felvennünk annak helyére. Ez a függvény Pascal programozási nyelven az alábbi módon implementálható:

```
Function szamol(aktmut:lmmut):Byte;
Var ujszam:Byte;

Begin
  ujszam:=0;
  If aktmut^.adat.elozo<aktmut^.adat.kov Then
    felso:=aktmut^.adat.elozo+8
  Else
    felso:=aktmut^.adat.elozo;
  For i:=felso-1 Downto aktmut^.adat.kov+1 Do
    If Not Odd(i) Then
      inc(ujszam);
  szamol:=ujszam;
End;
```

□

### 3.3. Problémák a vázkijelölés során

Ebben a fejezetben azon problémákat vizsgáljuk, amelyek a váz kijelölése során lépnek fel. Az ezek megoldását szolgáló elméleti eredményeinket pedig a következő fejezetekben ismertetjük. Elsőként tekintsük át, hogy milyen problémák adódhatnak a kontúr-reprezentáció használatából.

A törlés hatására előfordulhat, hogy korábban diszjunkt kontúrok részben, vagy teljes egészükben összeolvadnak. Ez nem jelent különösebb problémát, mert az általunk elkészített algoritmusban a közös kontúrponthoz minden olyan kontúrt reprezentáló listában tárolásra kerülnek, amelyhez tartoznak. Ez azt jelenti, ha például két kontúr egybeesik a törlések hatására, akkor a szóbanforgó kontúrok minden kontúrponthoz két kontúrt reprezentáló listában is tárolásra kerül. Az ilyen extrém esetek természetesen egy egyszerű utófeldolgozással kiszűrhetők.

Előfordulhat, hogy egy pixel hosszúságú kontúrt kapunk az algoritmus végrehajtása során. Ezt úgy fogjuk tárolni a memóriában, hogy felépítünk egy egyelemű listát. Ennek első (és egyetlen) elemének  $\text{kov}_{\text{lm}}$  mezője önmagára mutat. Ezzel szemben állományban csak a pont koordinátáit tároljuk, hiszen azt, hogy egy pont önmagának a rákövetkezője Freeman-féle lánczóddal nem is lehet leírni. Ezeket a speciális kontúrokat a további vizsgálatokból kizárjuk, mivel az őket reprezentáló listában szereplő  $\text{elozo}$ , illetve  $\text{kov}$  mezők értékei nem a valós helyzetet írják le. Ez a kizárás a vázkijelölés szempontjából is jogos, hiszen izolált pontok vázai önmaguk az izolált pontok, így az ezeket reprezentáló listákkal nem kell foglalkozni a későbbiekben.

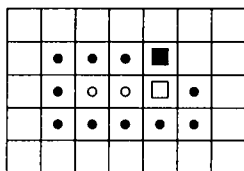
Az algoritmus végrehajtása során előfordulhat, hogy a kontúr kezdőpontjának megfelelő  $P$  pontot töröljük. Ebben az esetben a kontúrnak új kezdőpontja lesz. Amennyiben a törölt  $P$  kontúrponthoz új objektumpontokat veszünk fel, akkor az újonnan elsőként felvett objektumpont lesz a kontúr új kezdőpontja. A 3.4. Tétel alapján előfordulhat, hogy a törölt  $P$  kontúrponthoz nem veszünk fel új objektumpontot. Ebben az esetben a kontúr kezdőpontjának rákövetkezője lesz a kontúr új kezdőpontja.

A 10. ábra azt az esetet illusztrálja, amikor a kontúr kezdőpontjának törlése után új objektumpontokat veszünk fel a kontúrba, amíg a 11. ábra azt az esetet illusztrálja, amikor nem veszünk fel új objektumpontokat. A fent említett ábrákon a  $\bullet$  jelöli a kontúrponthoz,  $\circ$  a belső pontokat,  $\blacksquare$  a kontúr kezdőpontját, amíg  $\square$  a kontúr új kezdőpontját a régi törlés után.

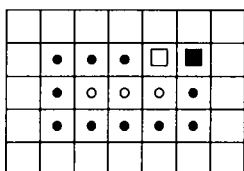
Bizonyos esetekben előfordulhat, hogy több kontúrponthoz törlésekor is ugyanazt az objektumpontot kell felvennünk. Ez hibát nem okoz, hiszen a következő iterációban ez a kontúrponthoz olyan tulajdonságú lesz, amelyen a kontúrkövetés során többször is áthaladunk.

Előfordulhat azonban az is, hogy bizonyos esetekben egy adott iterációban az utoljára felvett objektumpont ugyanaz lesz, mint az adott iterációban elsőként felvett objektumpont. Ez szemléletesen azt jelenti, hogy a kontúr kezdőpontját a kontúrt reprezentáló lista végére is felvesszük. Ez viszont már jelentős hibát eredményezhet, hiszen a listában egy pontnak nem lehet önmaga a rákövetkezője, il-





10. ábra. A kontúr új kezdőpontjának kijelölése, ha a régi törlése után új objektumpontokkal bővítjük a kontúrt



11. ábra. A kontúr új kezdőpontjának kijelölése, ha a régi törlése után új objektumpontokkal nem bővítjük a kontúrt.

letve megelőzője (kivéve az egyetlen pixelből álló kontúrokat reprezentáló listákat, amiket viszont külön kezelünk). Egyszerűen kiküszöbölhető ez a hiba is, hiszen elegendő az iteráció végén ellenőrizni, hogy az új kezdőpont abszolút koordinátái, illetve az adott kontúrt reprezentáló listabeli megelőzőjének abszolút koordinátái megegyeznek-e, és ha igen, akkor törölni kell a listából a kezdőpont megelőzőjét. A 10. és 11. ábra bármelyikén ez a helyzet fog előállni az iteráció végén.

Egy kontúrpon törlése során a  $C$ -szomszédjainak adatai, pontosabban a rákövetkezőjének előző, illetve megelőzőjének következője meg fognak változni. Ezen mezők új adatainak meghatározását a következő alfejezetben fogjuk tárgyalni.

### 3.4. A törölt pont $C$ -szomszédainak új adatai

Mint azt az előző fejezetben már említettük, egy  $P$  kontúrpon törlése során a  $C$ -szomszédaihoz tartozó adatokban változás áll be. Mielőtt megadnánk az új adatok meghatározásának módját egy lemmát bizonyítunk, amely a későbbiekben nagy segítségünkre lesz, és bizonyítása egyszerű számítással belátható.

3.5. LEMMA. Legyenek  $A$  és  $B$  egymással 8-szomszédos pontok. Ha  $\gamma_B(A) = i$ , akkor a  $\gamma_A(B) = (i + 4) \bmod 8$ .

3.6. TÉTEL. Legyen  $L$  egy kontúr,  $P$  egy törölni kívánt pontja,  $E$  és  $K$   $P$ -t megelőző, illetve követő  $C$ -szomszédja. Ekkor függetlenül attól, hogy  $P$  törlése után az  $L$  kontúrba hány belső pontot veszünk fel, a törölt  $P$  pont  $C$ -szomszédainak adataiban bekövetkező változás a következőképpen jellemezhető:

$$\gamma_E^{-1}(i + 1) = EK,$$

$$\gamma_K^{-1}(j - 1) = KE,$$

ahol  $i = \gamma_E(P)$ ,  $j = \gamma_K(P)$ ,  $EK$  az  $E$  kontúrpoint új rákövetkezője és  $KE$  a  $K$  kontúrpoint új megelőzője. A fentiek teljesülnek a  $\gamma_P(K) = (\gamma_P(E) - 1) \bmod 8$  speciális eset kivételével. Ebben az esetben, ha  $E$  4-szomszédja  $P$ -nek, akkor

$$\gamma_E^{-1}(i + 2) = EK,$$

$$\gamma_K^{-1}(j - 1) = KE.$$

Ha pedig  $K$  4-szomszédja  $P$ -nek, akkor

$$\gamma_E^{-1}(i + 1) = EK,$$

$$\gamma_K^{-1}(j - 2) = KE.$$

*Bizonyítás.* Legyen  $i = \gamma_E(P)$  és  $j = \gamma_K(P)$ . Ekkor a 3.5. Lemma alapján  $\gamma_P(E) = (i + 4) \bmod 8$  és  $\gamma_P(K) = (j + 4) \bmod 8$ . Ha  $i$  páros (páratlan), akkor értelemszerűen  $(i + 4) \bmod 8$  is az. Hasonló állítás teljesül a  $j$  esetén is.

Most tekintsük azt az esetet, amikor a  $j$  értéke páros. Ekkor a  $K$  és a  $P$  4-szomszédosak. Ebben az esetben teljesülnek a következő egyenlőségek:

$$(1) \quad \gamma_K^{-1}((j + 1) \bmod 8) = \gamma_P^{-1}((j + 4 - 2) \bmod 8)$$

$$(2) \quad \gamma_K^{-1}((j + 2) \bmod 8) = \gamma_P^{-1}((j + 4 - 1) \bmod 8)$$

$$(3) \quad \gamma_K^{-1}((j - 1) \bmod 8) = \gamma_P^{-1}((j + 4 + 2) \bmod 8)$$

$$(4) \quad \gamma_K^{-1}((j - 2) \bmod 8) = \gamma_P^{-1}((j + 4 + 1) \bmod 8)$$

Ezek az egyenlőségek négy pontot reprezentálnak, amelyek a  $K$  és  $P$  pontok 8-szomszédosságának a metszetét képezik.

Könnnyen látható, hogy a  $P$  pont belső oldali szomszédjai a  $\gamma_P^{-1}((j + 4 + 1) \bmod 8)$  ponttól a  $\gamma_P^{-1}((i + 4 - 1) \bmod 8)$  pontig található tartományból kerülnek ki. Ezek alapján az (1)–(2) egyenlőségek által jellemzett pontok külső oldali pontok lesznek.

A  $P$  kontúrpoint törlése után a  $K$  pont új megelőzője ( $KE$ ) az az objektum-pont lesz, ami  $P$ -nek belső oldali 4-szomszédja,  $K$ -nak pedig 8-szomszédja. Ez pedig csak a  $\gamma_P^{-1}((j + 4 + 2) \bmod 8)$  lehet, ami egybeesik  $\gamma_K^{-1}((j - 1) \bmod 8)$ -al, ahogy azt korábban láttuk. Az  $E$ -re analóg módon bebizonyítható, hogy az új rákövetkezője  $EK = \gamma_E^{-1}((i + 1) \bmod 8)$  lesz.

Tekintsük most azt az esetet, amikor  $j$  értéke páratlan. Ekkor  $K$  és  $P$  közvetett szomszédok. Ebben az esetben teljesülnek a következő egyenlőségek:

$$(5) \quad \gamma_K^{-1}((j + 1) \bmod 8) = \gamma_P^{-1}((j + 4 - 1) \bmod 8)$$

$$(6) \quad \gamma_K^{-1}((j - 1) \bmod 8) = \gamma_P^{-1}((j + 4 + 1) \bmod 8)$$

Az (5)–(6) egyenletek által reprezentált pontok alkotják a  $K$  és a  $P$  pontok 8-szomszédságának metszetét. Az (5) egyenlet által meghatározott pont külső oldali lesz.

A  $j$  páros esethez hasonlóan most is beláthatjuk, hogy a  $P$  törlése után  $K$  új megelőzője ( $KE$ ) a  $\gamma_P^{-1}((j+4+1) \bmod 8)$  lesz, ami egybeesik a  $\gamma_K^{-1}((j-1) \bmod 8)$ -cal, ahogy ezt korábban láttuk. Az  $E$ -re analóg módon látható, hogy az új rákövetkezője  $EK = \gamma_E^{-1}((i+1) \bmod 8)$  lesz.

A fent tárgyalt két eset nem fedi le azokat a konfigurációkat, amikor  $\gamma_P(K) = (\gamma_P(E) - 1) \bmod 8$  egyenlőség teljesül. A 12. ábrán látható konfigurációk, illetve ezek  $k \cdot 90^\circ$ -os ( $k \in \mathbb{Z}$ ) elforgatottjai írják le azokat az eseteket, amikor a fent említett tulajdonság teljesül.



12. ábra. A  $\gamma_P(K) = (\gamma_P(E) - 1) \bmod 8$  egyenlőséget teljesítő konfigurációk

Tekintsük az első esetet. Ha  $\gamma_E(P) = i$ , akkor  $\gamma_P(E) = (i+4) \bmod 8$  a 3.5. Lemma alapján. Hasonlóképpen, ha  $\gamma_P(K) = (i+4-1) \bmod 8$ , amiből  $\gamma_K(P) = (i+4-1+4) \bmod 8$ , azaz  $\gamma_K(P) = (i-1) \bmod 8$  következik.

Válasszuk meg a  $P$  törlése után az  $E EK$  rákövetkezőjét úgy, hogy  $\gamma_E(EK) = (i+1) \bmod 8$  teljesüljön. Majd válasszuk meg a  $K KE$  megelőzőjét úgy, hogy  $\gamma_K(KE) = (i-1-2) \bmod 8$  teljesüljön. Egyszerű számolással és a 3.5. Lemma alkalmazásával azt kapjuk, hogy az  $EK = K$  és  $KE = E$  egyenlőségek teljesülnek. A második eset hasonló módon bizonyítható, csak az  $E$  és  $K$  szerepe fog szimmetrikusan megváltozni.  $\square$

Megjegyezzük, hogy a tételben szereplő kivételes konfiguráció csak a törlés során alakulhat ki, hiszen egyébként a  $P$  pont nem lenne törölhető a 3.1. Definíció következtében.

### 3.5. A kontúrba bekerülő pontok adatainak meghatározása

A 3.4. alfejezetben ismertetett eredmények lehetővé teszik, hogy a kontúr tetszőleges  $P$  kontúrponkjának törlése után  $C$ -szomszédainak adatait meghatározzuk. Ez abban az esetben, ha a törölt  $P$  kontúrponthelyére nem kell új objektumpontokat felvennünk egyben ugrást jelent a kontúr következő kontúrponkjának vizsgálatára. Ebben a fejezetben azt fogjuk megvizsgálni, hogy ha a  $P$  kontúrponthelyére új objektumpontokat kell felvennünk annak helyére, akkor azokat milyen módon kell felvennünk a listába. Ehhez a következő megfontolásokat kell figyelembe venni.

**3.7. TÉTEL.** Legyen  $L$  egy kontúr, és  $P$  ennek egy törlendő pontja  $K$  és  $E$   $C$ -szomszédokkal. A kontúrkövetés sorrendjében az  $E$  legyen a  $P$  megelőző,  $K$  pedig

következő  $C$ -szomszédja a 3.6. Tétel alapján megváltoztatott adatokkal. Ha a  $P$  törlése után egyetlen  $A$  pontot vehetünk fel a kontúrba, akkor az  $L'$  kontúrban az  $E$  pont rákövetkezője és a  $K$  pont megelőzője az  $A$  pont lesz. Az  $A$  pont adatait a következőképpen határozhatjuk meg:

$$\begin{aligned} A.\text{elozo} &= (E.\text{kov} + 4) \bmod 8, \\ A.\text{kov} &= (K.\text{elozo} + 4) \bmod 8. \end{aligned}$$

*Bizonyítás.* A 3.5. Lemma segítségével egyszerű számítás mutatja.  $\square$

3.8. TÉTEL. Legyen  $L$  egy kontúr, és  $P$  ennek egy törlendő pontja  $K$  és  $E$   $C$ -szomszédokkal. A kontúrkövetés sorrendjében az  $E$  legyen a  $P$  megelőző,  $K$  pedig következő  $C$ -szomszédja a 3.6. Tétel alapján megváltoztatott adatokkal. Ha a  $P$  törlése után két pontot,  $A$ -t és  $B$ -t vehetjük fel a kontúrba, akkor az  $L'$  kontúrban az  $E$  pont rákövetkezője az  $A$  pont, az  $A$  pont megelőzője az  $E$  pont, rákövetkezője a  $B$  pont, a  $B$  pont megelőzője az  $A$  pont, a rákövetkezője a  $K$  pont, a  $K$  pont megelőzője a  $B$  pont lesz. Az  $A$  és a  $B$  pont adatait a következőképpen határozhatjuk meg:

$$\begin{aligned} A.\text{elozo} &= (E.\text{kov} + 4) \bmod 8, \\ A.\text{kov} &= (a + 5) \bmod 8, \\ B.\text{elozo} &= (a + 1) \bmod 8, \\ B.\text{kov} &= (K.\text{elozo} + 4) \bmod 8, \end{aligned}$$

ahol  $a = \gamma_P(A)$ .

*Bizonyítás.* Az  $A.\text{elozo}$  és a  $B.\text{kov}$  értékeit a 3.5. Lemma és a 3.7. Tétel alapján számíthatjuk ki.

Arra való tekintettel, hogy az  $A$  pont  $P$ -től  $a$  irányban van, az alábbi egyenlőségek teljesülnek:

$$\begin{aligned} \gamma_P^{-1}(a) &= A, \\ \gamma_A^{-1}((a + 4) \bmod 8) &= P. \end{aligned}$$

A 3.5. Lemma és a  $P$  pont és az  $A$  pont 4-szomszédos viszonyának következtében teljesülnek a következő egyenlőségek is:

$$\begin{aligned} \gamma_P^{-1}((a - 1) \bmod 8) &= \gamma_A^{-1}((a + 4 + 2) \bmod 8), \\ \gamma_P^{-1}((a - 2) \bmod 8) &= \gamma_A^{-1}((a + 4 + 1) \bmod 8). \end{aligned}$$

Mivel  $A$ -t és  $B$ -t a tétel állításában meghatározott sorrendben szűrtük be és az  $A$  pont és a  $B$  pont 8-szomszédos egymással,  $P$ -nek pedig 4-szomszédjai, így

$$B = \gamma_P^{-1}((a - 2) \bmod 8) = \gamma_A^{-1}((a + 4 + 1) \bmod 8)$$

azaz

$$A.kov := ((a + 5) \bmod 8)$$

Ez a 3.5. Lemma alapján a következő egyenlőséget eredményezi:

$$B.elozo := (a + 5 + 4) \bmod 8 = (a + 1) \bmod 8.$$

Ezzel a tétel bizonyítását befejeztük.  $\square$

**3.9. TÉTEL.** Legyen  $L$  egy kontúr, és  $P$  ennek egy törlendő pontja  $K$  és  $E$   $C$ -szomszédokkal. A kontúrkövetés sorrendjében az  $E$  legyen a  $P$  megelőző,  $K$  pedig következő  $C$ -szomszédja a 3.6. Tétel alapján megváltoztatott adatokkal. Ha a  $P$  törlése után három pontot, nevezetesen  $A$ -t,  $B$ -t és  $D$ -t vehetjük fel a kontúrba, akkor az  $L'$  kontúrban a megnevezett pontok – a kontúrkövetés sorrendjében – a következő sorrendben lesznek  $E$ ,  $A$ ,  $B$ ,  $D$ ,  $K$ . A felsorolás egyben meghatározza minden pont esetén annak megelőző, illetve rákövetkező  $C$ -szomszédját. A megváltozott, illetve az új adatokat a következőképpen határozhatjuk meg:

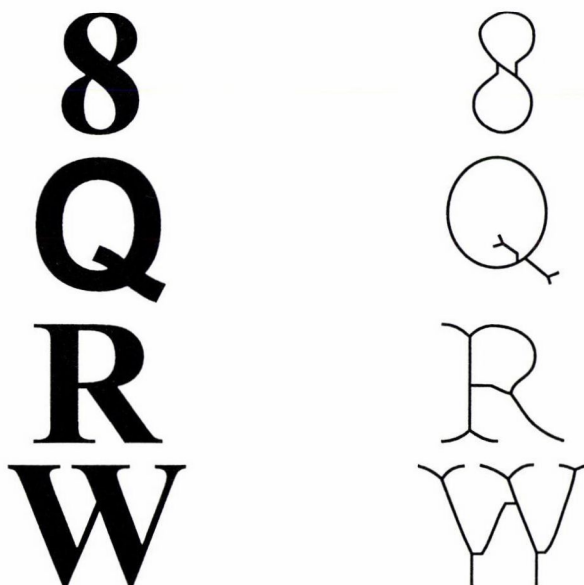
$$\begin{aligned} A.elozo &:= (E.kov + 4) \bmod 8 \\ A.kov &:= (P.elozo + 4) \bmod 8 \\ B.elozo &:= P.elozo \\ B.kov &:= P.kov \\ D.elozo &:= (P.kov + 4) \bmod 8 \\ D.kov &:= (K.elozo + 4) \bmod 8 \end{aligned}$$

*Bizonyítás.* A 3.8. Tételhez hasonlóan bizonyítható.  $\square$

## Összefoglalás

Cikkünkben egy speciális vázkijelölő algoritmust ismertettünk, amely úgynevezett kontúr-reprezentációjú bináris képeken üzemel. Részletesen közöltük azon elméleti eredményeinket is, amelyek lehetővé tették az algoritmus elméleti modelljének kidolgozását, illetve az algoritmus korrekt működésének garantálását. Ezen túl az algoritmus hatékony implementálásához szükséges elvi és gyakorlati megfontolások is megismerhetők a cikkből. A cikkben nem vizsgáltuk az algoritmusunk által szolgáltatott váz tulajdonságait, hiszen ez az algoritmus – ahogy már korábban jeleztük – Arcelli [1]-ben publikált algoritmusának kontúr-reprezentációra történő adaptációja. A fentiek következtében Arcelli korábbi vizsgálatainak eredményei változatlanul érvényesek az algoritmusunk által szolgáltatott vázra. A 13. ábrán illusztrációként néhány tesztábrát és azok – algoritmusunk által kinyert – vázait tekinthetjük meg.

Az algoritmusunk előnye az, hogy a kontúr-reprezentáció által biztosított kisebb memóriaigény lehetővé teszi nagyobb méretű bináris képek vázának hatékony kinyerését. Természetesen a kisebb memóriaigény ellensúlyaként a kontúr-reprezentáció kezeléséből eredő nagyobb számítási igénnyel kell számolnunk.



13. ábra. Objektumok és azok vázai

### Irodalomjegyzék

- [1] C. Arcelli, Pattern Thinning by Contour Tracing, *Computer Graphics and Image Processing*, **17** (1981), 130–144.
- [2] L. A. Breene, Quadtrees and Hypercubes: Grid Embedding Strategies Based on Spatial Data Structure Addressing, *The Computer Journal*, **36** (1993), 562–569.
- [3] A. Fazekas, G.T. Herman and S. Matej, On Processing Binary Pictures Via Their Projections, *The International Journal of Imaging Systems & Technology*, **9** (1998), 99–100.
- [4] L. Lam, S.-W. Lee and C. Y. Suen, Thinning Methodologies – A Comprehensive Survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14** (1992), 869–882.
- [5] J. Piper, Interval Skeletons, *11th IAPR*, **III** (1992), 468–471.
- [6] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, Academic Press (1992).
- [7] J. Serra, Introduction to Mathematical Morphology, *Computer Vision, Graphics, and Image Processing*, **35** (1986), 283–305.

(Beérkezett: 1999. november 27.)

FAZEKAS ATTILA  
DEBRECENI EGYETEM MATEMATIKAI ÉS INFORMATIKAI INTÉZET,  
H-4010 DEBRECEN  
PF. 12  
*E-mail:* FATTILA@MATH.KLTE.HU

SÁNTA ISTVÁN  
DEBRECENI EGYETEM MATEMATIKAI ÉS INFORMATIKAI INTÉZET,  
H-4010 DEBRECEN  
PF. 12  
*E-mail:* SANTAI@DRAGON.KLTE.HU

## SKELETONIZATION ON CONTOUR-REPRESENTATION

ATTILA FAZEKAS AND ISTVÁN SÁNTA

In digital image processing, skeletonization is a frequently used technique, when line drawings are processed. In the literature, most of the procedures operate on pixel-pixel type images. In this paper, our aim is to give an algorithm, which operates on contour-representing binary images. This algorithm is an adaptation to contour-representation of Arcelli's classical algorithm based on contour analyses.





## LEÍRÓ LOGIKÁK AZ ISMERETÁBRÁZOLÁSBAN

BOGNÁR KATALIN\*

Debrecen

Jelen dolgozatban a leíró logikákat mutatjuk be. Ebben az ismeretábrázolási technikában koncepciók, szerepek és individuumok jelenítik meg a szakterület ismeretanyagát, lehetővé téve a hierarchikus szerkezetkialakítást, és következtetési módszerként az osztályozás és egyedestítés alkalmazását. A leíró logikák leírását az [1] és [2] dolgozatok részletesen megadják. Ebben a dolgozatban két példán keresztül világítjuk meg a leíró logikák használatát a tudásbázisok létrehozásában és az ismeretek következtetéssel történő bővítésében. Összehasonlítjuk az objektum alapú ismeretábrázolást és a leíró logikákat.

Kulcsszavak: leíró logikák, ismeretábrázolás, következtetés, osztályozás, alárendelés.

### 1. Bevezetés

Az ismeretábrázolás a 80-as évek óta egyik központi problémája a mesterséges intelligencia kutatásoknak. A szakértő rendszerek felépítésében alapvető szerepet játszanak a tudásbázis és a következtetőrendszerek, felváltva a hagyományos adat-szerkezet és algoritmus fogalmakat. Egy tudásbázison alapuló rendszer olyan program, amely lehetővé teszi, hogy alkalmazási területén következtetéseket vonjunk le bizonyos problémák megoldására. Számos kutatás folyik, amely az ismeretek ábrázolására és a következtetési módszerek vizsgálatára irányul. Ezek közé tartozik a leíró logikáknak (description logics, terminological logics, concept languages) nevezett irányzat, amely az elsőrendű logika, a szemantikus hálók és a frame alapú nyelvek tanulmányozása, felhasználása során alakult ki. A leíró logikák tulajdonképpen egy ismeretábrázolási nyelvcsaládot alkotnak. A formalizmusukban a koncepció, individuum és szerep fogalmak jelennek meg. A koncepció individuumok halmazának reprezentálására szolgál, míg a szerep az individuumok közötti biná-

---

\* Jelen publikáció a magyar-francia kormányközi tudományos és technológiai együttműködés keretében, az OMFb és külföldi szerződéses partnere, Le Ministère des Affaires Étrangères által támogatott kutatási együttműködés (Balaton F-4/98) eredményeképpen jött létre.

ris relációt ábrázolja. A koncepció az alkalmazási terület egy általános, generáló egysége, az individuum speciális, a koncepció megjelenési formája, annak tulajdonságait viseli. A koncepció, szerep és individuum a következő alapelveknek felelnek meg:

- A koncepció és a szerep strukturális leírásában konstruktorok vesznek részt. A koncepció és a szerep leírásához egy szemantika kapcsolódik az interpretáción keresztül. A különböző műveleteket ezen szemantikával összhangban hajtjuk végre.
- Az ismereteket különböző szinteken vesszük figyelembe. A koncepciók, szerepek ábrázolása és műveleteik a terminológia szintjén, az individuumok leírása és műveleteik a tények és a hozzárendelések szintjén jelennek meg. A szakirodalomban a terminológia szintjét TBox-nak, a tények és a hozzárendelések szintjét ABox-nak nevezzük.
- A koncepciókat (és esetenként a szerepeket) hierarchiába rendezhetjük a rajtuk értelmezett alárendelés reláció alapján. Azt mondhatjuk, hogy egy  $C$  koncepció alárendeli a  $D$  koncepciót, ha  $C$  általánosabb, mint  $D$  abban az értelemben, hogy a  $D$  által reprezentált individuumok halmazát  $C$  tartalmazza.
- A leíró logikák következtetőrendszerében két művelet jelenik meg: a klasszifikáció (classification) és az egyedesítés (instantiation). A klasszifikációt a koncepciókra és a szerepekre alkalmazzuk. Lehetővé teszi, hogy egy adott koncepció, vagy szerep helyét meghatározzuk a hierarchiában. Az egyedesítés lehetővé teszi, hogy megtaláljuk azt a koncepciót, amelynek egy adott individuum a megjelenési formája lehet. Ez a fogalom eltér az objektum-orientált nyelvekben szokásos egyedesítés fogalmától, hiszen ott egy adott osztályból hozunk létre egyedekeket.

## 2. A leíró nyelvek szintaxisa és szemantikája

### 2.1. A leíró nyelvek szintaxisa

Koncepción az ábrázolandó világ elemeit jelentő individuumok halmazát értjük. A koncepciók halmazát  $\Delta$ -val jelöljük. Szerepen az individuumok közötti bináris relációt értjük.

Leíró nyelvnek nevezzük a

(koncepció-nevek, individuum-nevek, szerep-nevek, konstruktorok)

négyest, ahol a koncepció-nevek különböző koncepciókat, az individuum-nevek individuumokat, a szerep-nevek szerepeket szimbolizálnak. A konstruktorok a következők lehetnek: konjunkció ( $\sqcap$ ), diszjunkció ( $\sqcup$ ), negáció ( $\neg$ ), univerzális kvantor ( $\forall$ ), egzisztenciális kvantor ( $\exists$ ), számosság ( $\geq n$ ,  $\leq n$ ).

Az egyes konstruktorok a megfelelő definíció szerint koncepció- és szerepneveket kötnek össze, és így koncepció- és szerep-kifejezések jönnek létre. A koncepciónevek önmagukban koncepció-kifejezések. Ha  $C$  és  $D$  koncepció-kifejezés, akkor  $C * D$  és  $\star C$  is koncepció-kifejezések, ahol  $*$  valamely bináris,  $\star$  valamely unáris konstruktor. A továbbiakban a koncepció-neveket  $A, B$ , a szerep-neveket  $P$ , az individuumok nevét  $a, b, o$ , a koncepció-kifejezéseket  $C, D$ , a szerep-kifejezéseket  $Q, R$  jelöli. A top ( $\top$ ) és bottom ( $\perp$ ) speciális koncepciók; a top a legáltalánosabb, míg a bottom a leginkább specifikus koncepciót jelöli.

A különböző leíró nyelveket a megengedett konstruktorok határozzák meg.

Az alapszavak az  $\mathcal{FL}$ , amely konjunkció, univerzális kvantor és a nem minősített egzisztenciális kvantor konstruktorokat tartalmaz. Ezen alapul a legáltalánosabban vizsgált  $\mathcal{AL}$  nyelv, amely az előbbieken kívül (azaz az  $\mathcal{FL}$ -ből származtatva) tartalmazza a top, bottom koncepciókat, valamint a koncepció-név negációt (azaz a koncepció-név negálható, de koncepció-kifejezés nem).  $\mathcal{AL} = \{\top, \perp, \neg A, C \sqcap D, \forall R.C, \exists R\}$ . Az  $\mathcal{AL}$  nyelvcsaládot a megengedett konstruktorokkal kiegészítve kapjuk az  $\mathcal{AL}[\mathcal{U}][\mathcal{C}][\mathcal{E}][\mathcal{N}][\mathcal{R}]$  nyelveket, ahol  $\mathcal{U}$  a diszjunkció,  $\mathcal{C}$  a negáció,  $\mathcal{E}$  az egzisztenciális kvantor,  $\mathcal{N}$  a számosság,  $\mathcal{R}$  a szerep konjunkció konstruktorokat jelöli.

## 2.2. A leíró nyelvek szemantikája

A koncepciót az interpretációs alaphalmaz részhalmazaként, míg a szerepet az alaphalmaz önmagával alkotott Descartes szorzatának részhalmazaként interpretáljuk.

- Az interpretációs alaphalmaz ( $\mathcal{O}$ ) rögzítésével az a individuum interpretációja  $a^I \in \mathcal{O}$ .
- A koncepció-név interpretációja  $A^I \subseteq \mathcal{O}$ .
- A  $C$  koncepció  $C^I$  interpretációja a  $C$  koncepciót alkotó individuumok interpretációiból álló halmaz, azaz ha  $C = \{c_i\}$ , akkor  $C^I = \{c_i^I\}$ , tehát  $C^I \subseteq \mathcal{O}$ .
- A  $\Delta^I$  az összes  $C^I$  halmaza, azaz az interpretációs alaphalmaz ( $\mathcal{O}$ ) hatványhalmaza.
- Az  $R$  szerep interpretációja  $R^I \subseteq \mathcal{O} \times \mathcal{O}$ .

## 2.3. Az $\mathcal{ALCCNR}$ nyelv szemantikája

**2.1. Definíció.** Egy  $\mathcal{I} = (\Delta^I, \cdot^I)$  interpretáció egy interpretációs alaphalmaz és egy interpretációs függvény együttese, ahol az  $\cdot^I$  interpretációs függvény egy koncepciót hozzárendel a  $\Delta^I$  egy részhalmazához és egy szerepet a  $\Delta^I \times \Delta^I$  egy részhalmazához úgy, hogy a következő azonosságok fennálljanak.

$$\begin{aligned}\top^I &= \Delta^I \\ \perp^I &= \emptyset \\ (C \sqcap D)^I &= C^I \cap D^I\end{aligned}$$

$$\begin{aligned}
(C \sqcup D)^I &= C^I \cup D^I \\
(\neg C)^I &= \Delta^I \setminus C^I \\
(\forall R.C)^I &= \{a \in \Delta^I \mid \forall b : (a, b) \in R^I \rightarrow b \in C^I\} \\
(\exists R.C)^I &= \{a \in \Delta^I \mid \exists b : (a, b) \in R^I \wedge b \in C^I\} \\
(\geq n R)^I &= \{a \in \Delta^I \mid \{b \in \Delta^I \mid (a, b) \in R^I\} \geq n\} \\
(\leq n R)^I &= \{a \in \Delta^I \mid \{b \in \Delta^I \mid (a, b) \in R^I\} \leq n\} \\
(R_1 \sqcap \dots \sqcap R_n)^I &= R_1^I \cap \dots \cap R_n^I
\end{aligned}$$

**2.2. Definíció.** Két koncepciót  $(C, D)$  ekvivalensnek nevezünk  $(C \equiv D)$ , ha  $C^I = D^I$  minden  $I$  interpretációban.

Az egzisztenciális kvantornak  $(\exists R.C)$  egy speciális esete a nem minősített egzisztenciális kvantor  $(\exists R)$ , amikor  $C \equiv \top$ . Interpretációja:  $(\exists R)^I = \{a \in \Delta^I \mid \exists b : (a, b) \in R^I\}$ .

Az alábbi táblázatok összefoglaló képet adnak az alapfogalmak és konstruktorok jelölés-rendszeréről és értelmezéséről.

Alapfogalmak	Szintaxis	Szemantika
Koncepció	$A$	$A^I \subseteq \Delta^I$
top	$\top$	$\Delta^I$
bottom	$\perp$	$\emptyset$
individuumok ( $\mathcal{O}$ )	$\{a_1, a_2, \dots, a_n\}$	$\{a_1^I, a_2^I, \dots, a_n^I\}$
szerep	$P$	$P^I \subseteq \Delta^I \times \Delta^I$

### 1. Alapfogalmak

Konstruktorok	Szintaxis	Szemantika
konjunkció	$C \sqcap D$	$C^I \cap D^I$
diszjunkció ( $\mathcal{U}$ )	$C \sqcup D$	$C^I \cup D^I$
negáció ( $\mathcal{C}$ )	$\neg C$	$\Delta^I \setminus C^I$
univerzális kvantor	$\forall R.C$	$\{a_1 \mid \forall a_2 : (a_1, a_2) \in R^I \rightarrow a_2 \in C^I\}$
egzisztenciális kvantor ( $\mathcal{E}$ )	$\exists R.C$	$\{a_1 \mid \exists a_2 : (a_1, a_2) \in R^I \wedge a_2 \in C^I\}$
nem minősített egzisztenciális kvantor	$\exists R$	$\{a_1 \mid \exists a_2 : (a_1, a_2) \in R^I \wedge a_2 \in \Delta^I\}$
számosság ( $\mathcal{N}$ )	$(\geq n R)$ $(\leq n R)$	$\{a_1 \mid \{a_2 \mid (a_1, a_2) \in R^I\} \geq n\}$ $\{a_1 \mid \{a_2 \mid (a_1, a_2) \in R^I\} \leq n\}$
Szerep konjunkció ( $\mathcal{R}$ )	$Q \sqcap R$	$Q^I \cap R^I$

### 1. Koncepció- és szerep-formáló konstruktorok

### 3. Hierarchia a koncepciók és a szerepek körében

**3.1. Definíció.** Egy  $C$  koncepció alárendeltje a  $D$  koncepciónak, (jelölésben:  $C \sqsubseteq D$ ), ha tetszőleges  $\mathcal{I}$  interpretáció esetén  $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ .

Az alárendelés reláció reflexív, tranzitív és antiszimmetrikus, tehát egy parciális rendezési reláció, amely a koncepciókat egy hierarchiába szervezi. Ebben a hierarchiában a koncepciókat egyrészt saját lokális leírójuk jellemzi, másrészt az alárendeltjeikkel megosztott leírásuk (mint ahogyan az objektum-orientált nyelvekben az alá- és fölérendelt osztályoknál szokásos). Az így kialakult hierarchiában van egy „maximális” elem a top koncepció, amelynek minden más koncepció alárendeltje, és egy „minimális” elem, a bottom, amely valamennyi koncepciónak alárendeltje.

Mivel  $\Delta^{\mathcal{I}}$  az  $\sqsubseteq$ -re nézve háló; a koncepciók konjunkciója és diszjunkciója tulajdonképpen halmaz metszet és únió, amelyekre teljesülnek a hálóaxiómák:

- $A \sqcap A \equiv A$  és  $A \sqcup A \equiv A$  (idempotencia)
- $A \sqcap B \equiv B \sqcap A$  és  $A \sqcup B \equiv B \sqcup A$  (kommutativitás)
- $A \sqcap (B \sqcap C) \equiv (A \sqcap B) \sqcap C$  és  
 $A \sqcup (B \sqcup C) \equiv (A \sqcup B) \sqcup C$  (asszociativitás)
- $A \sqcap (A \sqcup B) \equiv A$  és  $A \sqcup (A \sqcap B) \equiv A$  (elnyelés)

További tulajdonságok:

- Ha  $D \sqsubseteq C$  és  $D \sqsubseteq E$ , akkor  $D \sqsubseteq C \sqcap E$
- Ha  $D \sqsubseteq C$  és  $E \sqsubseteq C$ , akkor  $D \sqcup E \sqsubseteq C$
- Ha  $D \sqsubseteq C$ , akkor  $D \sqcap X \sqsubseteq C$ , ahol  $X$  tetszőleges koncepció
- Ha  $D \sqsubseteq C$ , akkor  $D \sqsubseteq C \sqcup X$ , ahol  $X$  tetszőleges koncepció.

Az  $\mathcal{ALCN}$  nyelv hálót alkot az alárendelés műveletét tekintve, ahol a  $C$  és  $D$  koncepciók legkisebb felső korlátja  $C \sqcap D$ , legnagyobb alsó korlátja  $C \sqcup D$ .

### 4. A leíró ismeretbázis fogalma

A leíró nyelvekben az ismeretábrázolás két szinten valósul meg. A terminológia szintjén vezetjük be a koncepciókat, a szerepeket és az adott  $\mathcal{ALCN}$  leíró nyelvnek megfelelően az alárendelési relációkat. A koncepciók és szerepek lehetnek primitívek (atomiak) vagy összetettek (definiáltak). A primitív koncepciókat (szerepeket) alárendelési relációval adjuk meg, az összetett koncepciókat (szerepeket) pedig konstruktorok segítségével (jelölésben:  $\doteq$ ). A tények és a hozzárendelések szintjén az egyes koncepciókhoz tartozó individuumokat és az egyes szerepekhez

tartozó individuum párokat mint tényeket sorolunk fel. Jelölésben a hozzárendelések  $C(a)$  és  $R(a, b)$  alakúak. A hozzárendeléseket általánosan  $\alpha$  hozzárendelésnek jelölöm a további definíciókban.

**4.1. Definíció.** Leíró tudásbázisnak nevezzük (jelölésben:  $\Sigma = (\mathcal{T}, \mathcal{A})$ ) az  $ALCN\mathcal{R}$  nyelvben a  $(\mathcal{T}, \mathcal{A})$  párost, ahol  $\mathcal{T}$  a koncepciók és szerepek leírása a nyelv eszközeivel,  $\mathcal{A}$  pedig a tények és egyed-hozzárendelések megadása  $C(a)$  vagy  $R(a, b)$  alakban.

**4.2. Definíció.** Azt mondjuk, hogy az  $\mathcal{I}$  interpretáció modellje a  $C$  koncepciónak, ha  $C^{\mathcal{I}} \neq \emptyset$ .

**4.3. Definíció.** Azt mondjuk, hogy egy  $C$  koncepció kielégíthető, ha létezik modellje.

**4.4. Definíció.** Legyen  $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$  egy interpretáció. A  $C(a)$  hozzárendelést kielégíti az  $\mathcal{I}$  interpretáció, ha  $a^{\mathcal{I}} \in C^{\mathcal{I}}$ ; az  $R(a, b)$  hozzárendelést kielégíti az  $\mathcal{I}$  interpretáció, ha  $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$ .

**4.5. Definíció.** Egy  $\mathcal{I}$  interpretáció modellje a  $\Sigma = (\mathcal{T}, \mathcal{A})$  leíró tudásbázisnak, ha  $\mathcal{I}$  kielégíti  $\mathcal{A}$  minden hozzárendelését.

**4.6. Definíció.** A  $\Sigma = (\mathcal{T}, \mathcal{A})$  leíró tudásbázis kielégíthető, ha létezik modellje.

**4.7. Definíció.** Az  $\alpha$  hozzárendelés logikai következménye a  $\Sigma = (\mathcal{T}, \mathcal{A})$  leíró tudásbázisnak, ha  $\Sigma$  minden modellje kielégíti  $\alpha$ -t. Jelölésben:  $\Sigma \models \alpha$ .

#### 4.1. Példa leíró tudásbázisra

Az alábbi példában a  $\mathcal{T}$  Tbox négy koncepciót vezet be.

- t1 Egy kurzus oktatója vagy professzor vagy egyetemi diplomával rendelkező diák (PhD hallgató).
- t2 A professzorok doktori diplomával rendelkező személyek.
- t3 Ha valakinek doktori diplomája van, akkor biztosan van egyetemi diplomája is.
- t4 A doktori és egyetemi diplomák különbözőek.

Az  $\mathcal{A}$  Abox hozzárendelések közül a2 azt mutatja, hogy János nem lehet professzor, hiszen legfeljebb egy diplomája van, s ez a1 és a3 miatt, azaz mert János tanítja a Prog-kurzust, feltétlenül egyetemi diploma.

Legyen  $\Sigma = (\mathcal{T}, \mathcal{A})$  ahol

$\mathcal{T} = \{$   
 $\text{SZEMÉLY} \sqsubseteq \top$   
 $\text{PROFESSZOR} \sqsubseteq \text{SZEMÉLY}$   
 $\text{DIÁK} \sqsubseteq \text{SZEMÉLY}$   
 $\text{KURZUS} \sqsubseteq \top$

FOKOZAT $\sqsubseteq \top$	
EGYETEMI $\sqsubseteq$ FOKOZAT	
DOKTORI $\sqsubseteq$ FOKOZAT	
tanító $\sqsubseteq$ toprole	
diploma $\sqsubseteq$ toprole	
$(\exists \text{tanító.KURZUS}) \sqsubseteq$	t1
$(\text{PROFESSZOR} \sqcup (\text{DIÁK} \sqcap (\exists \text{diploma.EGYETEMI})))$	
$\text{PROFESSZOR} \sqsubseteq (\exists \text{diploma.DOKTORI})$	t2
$(\exists \text{diploma.DOKTORI}) \sqsubseteq (\exists \text{diploma.EGYETEMI})$	t3
$(\text{DOKTORI} \sqcap \text{EGYETEMI}) \sqsubseteq \perp$	t4
}	
$\mathcal{A} = \{$	
tanító(János, Prog_kurzus)	a1
$(\leq 1 \text{ diploma})(\text{János})$	a2
KURZUS(Prog_kurzus)	a3
}	

A következő interpretáció egy modellje az előbbi  $\Sigma = (\mathcal{T}, \mathcal{A})$  leíró tudásbázisnak, ahol az interpretációs alaphalmaz

$$\mathcal{O} = \{\text{Jani}, \text{Programozás}, \text{Jani-egyetemi-diploma}\}.$$

Ekkor a

$$\text{János}^I = \text{Jani}$$

$$\text{Prog\_kurzus}^I = \text{Programozás}$$

$$\text{DIÁK}^I = \{\text{Jani}\}$$

$$\text{PROFESSZOR}^I = \emptyset$$

$$\text{KURZUS}^I = \{\text{Programozás}\}$$

$$\text{EGYETEMI}^I = \{\text{Jani-egyetemi-diploma}\}$$

$$\text{DOKTORI}^I = \emptyset$$

$$\text{tanító}^I = \{(\text{Jani}, \text{Programozás})\}$$

$$\text{diploma}^I = \{(\text{Jani}, \text{Jani-egyetemi-diploma})\}$$

interpretáció kielégíti  $\mathcal{A}$  minden hozzárendelését.

## 5. Következtetési eljárások egy leíró tudásbázisban

Egy leíró tudásbázisban az alábbi következtetési eljárások alkalmazhatók:

- Alárendelések ellenőrzése. Ezen eljárás segítségével eldönthetjük, hogy egy  $C$  koncepció alárendeli-e a  $D$  koncepciót vagy sem. Ez az alapja az osztályozási műveletnek, ami meghatározza egy koncepció közvetlen leszármazottait.
- Egy koncepció kielégíthetőségének ellenőrzése. Ennek során eldönthetjük, hogy egy koncepciónak létezik-e modellje, azaz vannak-e egyedei valamely interpretációban.

- Egy leíró tudásbázis kielégíthetőségének vizsgálata; itt ellenőrizzük, hogy létezik-e modellje.
- Egyedestítés. Ezen eljárás során ellenőrizzük, hogy egy  $b$  individuum egyede-e a  $C$  koncepciónak a  $\Sigma$  leíró tudásbázisban, azaz  $\Sigma \models C(b)$  teljesül-e. Pontosabban ez az eljárás azon koncepciókat keresi meg, amelyeknek a  $b$  individuum egyede és amelyek ugyanakkor a legspecifikusabbak az alárendelési hierarchiában.

### 5.1. Példa az egyedestítésre

A  $\mathcal{T}$  Tboxban primitív és definiált koncepciókra láthatunk példát. A személy és halmaz primitív koncepciók, amelyeket a  $\sqsubseteq$  jel vezet be, és a  $\top$  (top koncepció) mint a legáltalánosabb koncepció alárendeltjei. A  $\top$  tekinthető a koncepciók hierarchiájában a gyökérelemnek.

A  $\sqcap$  (konjunkció) konstruktor jelzi, hogy egy koncepciót több más koncepció konjunkciójaként hozunk létre, amelyek az így definiált koncepció közvetlen ősei. A  $\forall$  konstruktor a tag szerep érvényességi tartományát pontosítja. A  $\neg$  konstruktor negációt fejez ki, csak primitív koncepcióra alkalmazzuk. A  $\leq$  és  $\geq$  konstruktorok a tag szerep érvényességi tartományának számosságát korlátozzák.

Legyen

$$\mathcal{T} = \{$$

$$\text{SZEMÉLY} \sqsubseteq \top$$

$$\text{HALMAZ} \sqsubseteq \top$$

$$\text{FÉRFI} \sqsubseteq \text{SZEMÉLY}$$

$$\text{NŐ} \sqsubseteq (\text{SZEMÉLY} \sqcap (\neg \text{FÉRFI}))$$

$$\text{tag} \sqsubseteq \text{toprole}$$

$$\text{főnök} \sqsubseteq \text{tag}$$

$$\text{CSAPAT} \doteq (\text{HALMAZ} \sqcap (\forall \text{tag.SZEMÉLY}) \sqcap (\geq 2 \text{ tag}))$$

$$\text{KISCSAPAT} \doteq (\text{CSAPAT} \sqcap (\leq 5 \text{ tag}))$$

$$\text{MODERNCSAPAT} \doteq (\text{CSAPAT} \sqcap (\leq 4 \text{ tag}) \sqcap (\geq 1 \text{ főnök}) \sqcap (\forall \text{főnök.NŐ}))$$

$$\}$$

$$\mathcal{A} = \{$$

$$\text{MODERNCSAPAT}(\text{Trio})$$

$$\text{FÉRFI}(\text{Antal})$$

$$\text{SZEMÉLY}(\text{Erzsi})$$

$$\text{tag}(\text{Trio}, \text{Antal})$$

$$\text{tag}(\text{Trio}, \text{Péter})$$

$$\text{főnök}(\text{Trio}, \text{Erzsi})$$

$$(\leq 3 \text{ tag})(\text{Trio})$$

$$\}$$



- Definíáljon az  $(\leq n_D R_D)$  és  $(\leq n_C R_C)$  számosság-korlátozás két koncepciót. Bizonyítható alárendelési szabály: ha  $n_D \leq n_C$  és  $R_C \subseteq R_D$ , akkor  $(\leq n_D R_D) \subseteq (\leq n_C R_C)$ .  
Ebből adódik, hogy MODERNCSAPAT alárendeltje a KISCSAPATNAK. ( $4 \leq 5$  és  $R =$  tag mindkét esetben.)
- Bizonyítható alárendelési szabály: ha az  $o$  egyede a  $C$  koncepciónak és  $C$  alárendeltje a  $D$  koncepciónak, akkor  $o$  egyede a  $D$  koncepciónak is. Ebből adódik, hogy a Trio egyede a KISCSAPAT-nak, hiszen a Trio a MODERNCSAPAT koncepció egy egyede és a MODERNCSAPAT alárendeltje a KISCSAPATNAK.
- Bizonyítható alárendelési szabály: ha  $D_1(o)$  és  $D_2(o)$  hozzárendelések, valamint a  $C$  koncepciónak nem alárendeltje sem  $D_1$  sem  $D_2$ , de alárendeltje  $(D_1 \sqcap D_2)$ , akkor ebből következik a  $C(o)$  hozzárendelés.
- Bizonyítható alárendelési szabály, ha adva van egy szerep-hozzárendelés  $R(o, b)$  és  $o$  egy egyede egy  $(\forall R.C)$  alakú koncepciónak, akkor  $C(b)$  következik, azaz  $b$  egyede  $C$ -nek.  
Ebben a példában, Péter és Antal tag relációban vannak a Trio-val és Trio egyede a MODERNCSAPAT-nak. Mivel MODERNCSAPAT alárendeltje a CSAPAT-nak, következésképpen a Trio egyede a CSAPAT-nak. A CSAPAT  $(\forall \text{tag.SZEMÉLY})$  definíciójából adódik, hogy Péter és Antal egyede a SZEMÉLY koncepciónak.  
Hasonlóan levezethető, hogy Erzsí egyede a Nő koncepciónak a főnök(Trio, Erzsí) és MODERNCSAPAT(Trio) hozzárendelésekből, valamint a MODERNCSAPAT  $(\forall \text{főnök.NŐ})$  alakú definíciójából.

## 6. A leíró logika, a klasszikus logika és az objektum alapú ismeretábrázolás

A leíró logikák fő jellemzői a koncepciókat és szerepeket leíró nyelv, a nyelvhez kapcsolódó interpretáció, valamint az alárendelési reláció. Természetes módon adódik tehát a kapcsolat a klasszikus formális logikával, ami a formulákon, az azokhoz kapcsolódó interpretáción és a levezetési szabályokon alapul. A leíró logikák koncepcióit tekinthetjük unáris, a szerepeket bináris predikátumoknak az alárendelést pedig levezetési szabálynak. A párhuzamosságot szemlélteti, hogy a koncepciókifejezések és a hozzárendelések lényegében speciális elsőrendű logikai formulák.

Például legyen adott a

$$C \doteq \exists \text{gyereke.NŐ} \sqcap \forall \text{gyereke.SZEMÉLY}$$

koncepció. Ez megfeleltethető az alábbi formulának:

$$\phi(x) = \exists y(\text{gyereke}(x, y) \wedge \text{NŐ}(y)) \wedge \forall z(\text{gyereke}(x, z) \rightarrow \text{SZEMÉLY}(z))$$

A  $\phi(x)$  formula modellje egyben modellje a  $C$  koncepciónak és megfordítva.

Ugyanakkor a leíró logikák bizonyos hasonlóságot mutatnak az objektum alapú ismeretábrázolással is. Egy objektum alapú ismeretábrázolási rendszer egy állapotot és egy viselkedést bezáró osztályokat tekint ismeretnek. Az állapotot és a viselkedést funkcionális valamint leíró tulajdonságok határozzák meg.

A leíró tulajdonságokhoz (attribútumok) típust, értéket és démonokat rendelhetünk. A funkcionális tulajdonságok (metódusok) üzenetküldés útján aktiválható eljárások vagy függvények. Az osztályok öröklődési hierarchiába rendezettek. Az osztályok egyedesíthetőek, s az így keletkezett objektumokkal végezhetnek műveletet az alkalmazások. Az objektum alapú ismeretábrázolási rendszerekben a következtetés megvalósítható öröklődéssel (egy tulajdonság értéke az osztályok közötti öröklődési relációból vezethető le), osztályozással (egy osztálynak az öröklődési hierarchiába való beillesztése, vagy egy egyednek egy osztályhoz rendelése útján) vagy szűréssel (egy bizonyos szűrőnek megfelelő objektumok kiválasztása révén).

A koncepció fogalom összevethető az osztályfogalommal, a szerep az osztály attribútumaival. Ugyanakkor a leíró logikáknál teljesen hiányzik a procedurális jelleg, ami alapvetően megkülönbözteti őket az objektum alapú ismeretábrázolástól. Ebből a szempontból tekinthetjük az objektum alapú ismeretábrázolást procedurálisnak, míg a leíró logikákat deklaratívnak.

Egy másik összehasonlítási szempont lehet az osztályozás, amikor az objektum alapú ismeretábrázolás szabad kezét ad a programozónak egy osztály elhelyezésére az öröklődési hierarchiába. A leíró logikák esetén egy új koncepció definiálásakor az osztályozási eljárás megadja a koncepció elhelyezését a hierarchiába az alárendelési szabálynak megfelelően.

További különbség, hogy az objektum alapú ismeretábrázolás esetén az osztályhoz rendelt tulajdonságok leíró jellegűek, azaz szükséges, de nem elégséges feltételt jelentenek. Azaz, ha egy  $o$  objektum egy  $C$  osztály egyede, akkor ez azt jelenti, hogy ily módon definiáltuk, és rendelkezik a  $C$  osztály tulajdonságaival. Ellenben, ha  $o$  rendelkezik a  $C$  osztály tulajdonságaival, ebből nem következtethetünk arra, hogy  $o$  valóban egyede is  $C$ -nek. A leíró logikákban az osztályozási mechanizmus definíciós szemantikán nyugszik, ahol a hierarchiába rendezésnek szükséges feltételei egyben elégséges feltételek is.

Ugyanakkor a kivételkezelés teljes egészében hiányzik a leíró logikákból, hiszen az alárendelési reláció ezt nem engedi meg.

## Irodalomjegyzék

- [1] A. Napoli, Une introduction aux logiques de descriptions, *Rapport de recherche INRIA Lorraine* (1997).
- [2] F. M. Donini, M. Lenzerini, D. Nardi and A. Schaerf, Reasoning in Description Logics, in: *Principles of Knowledge Representation*, Ed. G. Brewka (CSLI Publications, Stanford, USA, 1996), 191–236.

(Beérkezett: 2000. január 12.)

BOGNÁR KATALIN  
DEBRECENI EGYETEM  
MATEMATIKAI ÉS INFORMATIKAI INTÉZET  
DEBRECEN, 4010 PF. 12]] *E-mail*: BOGNAR@MATH.KLTE.HU

## DESCRIPTION LOGICS IN KNOWLEDGE REPRESENTATION

KATALIN BOGNÁR

This paper presents the formalism of description logics in knowledge representation allowing to represent domain knowledge with concepts, roles (classes of individuals, relations between classes) and individuals. Concepts and roles are organized with hierarchies where terminological reasoning (based on classification and instantiation) operates. Examples are given in sections 4 and 5 for illustration.

Key-words: description logics, knowledge representation, reasoning, classification, subsumption.



# TÖBBVÁLTOZÓS DISZKRÉT FÜGGVÉNYEK FÉLOLDALAS APPROXIMÁCIÓJA POLINOMOKKAL

NAGY GERGELY ÉS PRÉKOPA ANDRÁS

Budapest

A dolgozatban a diszkrét, magasabb rendű konvex függvényekkel kapcsolatban bebizonyítunk egy tételt, mely lehetővé teszi azok féloldalas approximációját többdimenziós Lagrange interpolációs polinomokkal. A tétel a folytonos, magasabb rendű konvex függvények vonatkozásában is új eredmény, általánosítja azt az egyváltozós formulát, mely a függvény és a Lagrange polinom különbségére vonatkozó maradéktagot adja meg. Tételünkéből könnyen levezethetők olyan korlátok, melyek magasabb rendű konvex függvények várható értékeire vonatkoznak.

## 1. Diszkrét függvények magasabb rendű konvexitása

Diszkrét függvényen olyan függvényt értünk, melynek értelmezési tartománya véges. A dolgozatban olyan véges halmazokon értelmezzük diszkrét függvényeket, melyek a számegegyenes részhalmazai illetve, többváltozós esetben, ilyen halmazok Descartes szorzatai.

Az  $f(z)$ ,  $z \in \{z_0, \dots, z_n\}$  diszkrét függvény elsőrendű osztott differenciáit az alábbi módon jelöljük, illetve definiáljuk:

$$(1.1) \quad [z_{i_1}, z_{i_2}; f] = \frac{f(z_{i_1}) - f(z_{i_2})}{z_{i_1} - z_{i_2}}.$$

A  $k$  rendű osztott differenciákat, a szokásos, induktív módon értelmezzük (lásd pl. Jordan [2], Popoviciu [5], Prékopa [8]).

Az  $f$  diszkrét függvényt  $k$  rendben konvexnek nevezzük, ha valamennyi  $k$  rendű osztott differenciája nemnegatív. Az elsőrendű konvexitás a függvény monoton növekvő voltát, a másodrendű konvexitás a függvény konvexitását jelenti, a hagyományos értelemben.

A fenti definíció némileg eltér Popoviciu [5] definíciójától, ő ugyanis a  $k$  rendű konvexitást a  $k + 1$  rendű osztott differenciák nemnegativitásával értelmezte.

A többváltozós függvényeket a számegegyenes véges részhalmazainak Descartes szorzatainak értelmezzük. Az  $f(z)$ ,  $z = (z_1, \dots, z_s) \in Z = Z_1 \times \dots \times Z_s$ ,

$Z_j = \{z_{j0}, \dots, z_{jn_j}\}$ ,  $j = 1, \dots, s$  függvény  $(k_1, \dots, k_s)$  rendű osztott differenciáját a  $Z$  halmaz valamely

$$(1.2) \quad \begin{aligned} Z_{I_1 \dots I_s} &= \{z_{1i}, i \in I_1\} \times \dots \times \{z_{si}, i \in I_s\} = \\ &= Z_{1I_1} \times \dots \times Z_{sI_s}, \end{aligned}$$

részhalmazához rendeljük, ahol  $|I_j| = k_j + 1$ ,  $j = 1, \dots, s$ , és az alábbi módon értelmezzük. Vesszük a  $z_1$  változó szerinti  $k_1$  rendű osztott differenciát, majd a  $z_2$  szerinti  $k_2$  rendű osztott differenciát, és így tovább, végül a  $z_s$  szerinti  $k_s$  rendű osztott differenciát. E műveleteket tetszőleges sorrendben végrehajthatjuk, az eredmény mindig ugyanaz.

Az  $f$  függvény  $(k_1, \dots, k_s)$  rendű osztott differenciájának a jelölésére a

$$(1.3) \quad [z_{1i}, i \in I_1; \dots; z_{si}, i \in I_s; f]$$

szimbólumot használjuk. A  $k_1 + \dots + k_s$  összeget az osztott differencia teljes rendjének nevezzük.

Az osztott differenciákra érvényesek az alábbi állítások.

Ha  $f(z)$ ,  $z \in Z$  egyváltozós diszkrét függvény és  $V_1, V_2 \in Z$ ,  $V_1 \cap V_2 = \emptyset$ , akkor

$$[V_1; [V_2; f]] = [V_2; [V_1; f]] = [V_1 \cup V_2; f].$$

Ha  $f(z)$ ,  $z \in Z = Z_1 \times Z_2$  diszkrét függvény és  $z_1 \in Z_1$ ,  $z_2 \in Z_2$ ,  $V_1 \subset Z_1$ ,  $V_2 \subset Z_2$ , akkor

$$[V_1; [z_1; V_2; f]] = [V_2; [V_1; z_2; f]] = [V_1; V_2; f].$$

**1.1. Definíció.** Az  $f(z)$ ,  $z \in Z$  (többváltozós) függvényt  $(m_1, \dots, m_s)$  rendben konvex diszkrét függvénynek hívjuk, ha bármely  $\{z_{ji}, i \in I_j\}$ ,  $|I_j| = m_j + 1$ ,  $j = 1, \dots, s$  esetén fennáll az alábbi reláció:

$$(1.4) \quad [z_{1i}, i \in I_1; \dots; z_{si}, i \in I_s; f] \geq 0.$$

**1.2. Definíció.** Egy  $f(z)$ ,  $z \in Z$  (többváltozós) függvényt  $m$  rendű konvex függvénynek nevezünk, ha valamennyi  $m$  teljes rendű osztott differenciája nemnegatív.

Ha  $f(z)$ ,  $g(z)$ ,  $z \in Z$  ugyanolyan rendben konvex, akkor ez a tulajdonság igaz az  $f(z) + g(z)$ ,  $z \in Z$  összegre is. A szorzat esetét tekintve, a következő állítást kapjuk.

**1.1. TÉTEL.** Ha  $f(z) \geq 0$ ,  $g(z) \geq 0$ ,  $z \in Z$  bármilyen  $i$  rendben konvex,  $1 \leq i \leq m$ , akkor ugyanez igaz az  $f(z)g(z)$ ,  $z \in Z$  függvényre is.

**Bizonyítás.** A szorzat osztott differenciája hasonló szabállyal kapható meg, mint a szorzat deriváltja. Ebből a tényből az állítás már könnyen levezethető.  $\square$

A magasabb rendű konvexitásra adott definíciónkban csak a koordináta tengelyek irányában vett osztott differenciákat használtuk.

Előfordulhat, hogy egy függvénynek az összes második teljes rendű osztott differenciája nemnegatív, viszont valamely (a tengelyektől különböző) egyenes mentén nem konvex a függvény. Erre példa a következő.

Legyen  $Z_1 = Z_2 = \{0, 1, 2\}$ , és definiáljuk az  $f(z)$ ,  $z \in Z_1 \times Z_2$  függvényt az alábbi módon:

$$\begin{aligned} f(0, 0) &= 0, & f(1, 0) &= 1,2, & f(2, 0) &= 2,6, \\ f(0, 1) &= 0,4, & f(1, 1) &= 2, & f(2, 1) &= 3,6, \\ f(0, 2) &= 1, & f(1, 2) &= 2,8, & f(2, 2) &= 4,6. \end{aligned}$$

A függvény nem konvex a  $(0, 2)$ ,  $(1, 1)$ ,  $(2, 0)$  egyenes mentén, mivel

$$f(1, 1) = 2 > \frac{1 + 2,6}{8} = \frac{f(0, 2) + f(2, 0)}{8}.$$

Ha az  $f(z)$ ,  $z \in Z$  függvényt az  $\bar{f}(z)$ ,  $z \in \bar{Z}$ ,  $\bar{Z} = [z_{10}, z_{1n_1}] \times \cdots \times [z_{s0}, z_{sn_s}]$  folytonos függvényből származtatjuk oly módon, hogy  $f(z) = \bar{f}(z)$ ,  $z \in Z$ , és  $\bar{f}(z)$   $(k_1, \dots, k_s)$  rendű deriváltjai  $\bar{Z}$  belsejében folytonosak és nemnegatívak, akkor  $f(z)$ ,  $z \in Z$  összes  $(k_1, \dots, k_s)$  rendű osztott differenciája nemnegatív.

Egy  $f(z)$ ,  $z \in Z$ , diszkrét,  $m$  rendben konvex függvény esetén általában nem könnyű konstruálni egy olyan  $\bar{f}(z)$ ,  $z \in \bar{Z}$  folytonos, nemnegatív  $m$  rendű deriváltakkal rendelkező függvényt, mely a  $Z$  diszkrét halmazon megegyezik  $f(z)$  értékeivel. Ha azonban az  $f$  függvény értelmezési tartományát megszorítjuk, akkor bizonyos esetekben mégis tudunk adni ilyen konstrukciót. Ezt fejezi ki az alábbi

1.2. TÉTEL. *Definiáljuk a  $Z_I$  diszkrét halmazt a következő módon:*

$$(1.5) \quad Z_I = \{(z_{i_1}, \dots, z_{i_s}) \mid (i_1, \dots, i_s) \in I\},$$

ahol

$$(1.6) \quad I = \{(i_1, \dots, i_s) \mid i_1 + \cdots + i_s \leq m, 0 \leq i_j \leq n_j, j = 1, \dots, s\},$$

$$m \leq n_j, \quad j = 1, \dots, s.$$

Ekkor létezik pontosan egy olyan  $L_I(z)$  polinom, melyre

$$L_I(z) = f(z), \quad z \in Z_I,$$

és  $L_I(z)$   $(k_1, \dots, k_s)$  rendű deriváltjai megegyeznek az  $f$  függvény  $(k_1, \dots, k_s)$  rendű osztott differenciáival az alábbi halmazon:

$$\{(z_{1i_1}, \dots, z_{si_s}) \mid 0 \leq i_j \leq k_j, j = 1, \dots, s\}.$$

Az  $L_I(\mathbf{z})$  polinom a következő:

$$(1.7) \quad L_I(z_1, \dots, z_s) = \sum_{\substack{i_1 + \dots + i_s \leq m \\ 0 \leq i_j \leq n_j, j=1, \dots, s}} [z_{10}, \dots, z_{1i_1}; \dots; z_{s0}, \dots, z_{si_s}; f] \times \\ \times \prod_{j=1}^s \prod_{h=0}^{i_j-1} (z_j - z_{jh}),$$

ahol, definíció szerint,  $i_j = 0$  esetén  $\prod_{h=0}^{i_j-1} (z_j - z_{jh}) = 1$ .

*Bizonyítás.* Könnyen ellenőrizhető, hogy  $L_I(\mathbf{z})$  megfelelő deriváltjai eleget tesznek a kívánalmaknak. A polinom egyértelműségének bizonyítása megtalálható Prékopa [8] cikkében (362. old., Theorem 4.1).  $\square$

*1.1. Megjegyzés.* Az  $L_I(\mathbf{z})$  polinomot a  $Z_I$  ponthalmazhoz tartozó Lagrange polinom Newton-féle alakjának nevezzük.

A továbbiakban gyakran felhasználjuk az egyváltozós Lagrange interpoláció elméletéből jól ismert alábbi formulát:

$$(1.8) \quad f(z) - L(z) = [z_0, \dots, z_k, z; f] \prod_{j=0}^k (z - z_j),$$

ahol  $L(z)$  a  $z_0, \dots, z_k$  alappontokhoz tartozó Lagrange polinom,

$$(1.9) \quad L(z) = \sum_{i=0}^k f(z_i) \frac{(z - z_0) \cdots (z - z_{i-1})(z - z_{i+1}) \cdots (z - z_k)}{(z_i - z_0) \cdots (z_i - z_{i-1})(z_i - z_{i+1}) \cdots (z_i - z_k)}.$$

Az (1.8) formulát általában intervallumon értelmezett függvények esetében használják. A mi esetünkben azonban nem csak az alappontok halmaza, hanem az  $f$  függvény értelmezési tartománya is véges.

## 2. A többváltozós diszkrét momentum probléma és kapcsolata a többváltozós Lagrange interpolációval

A többváltozós diszkrét momentum probléma (TDMP), mint látni fogjuk, szoros kapcsolatban van a többváltozós Lagrange interpoláció elméletével. Ebből kifolyólag, egyrészt a Lagrange interpolációban elért eredmények jól hasznosíthatók a momentum problémák terén, másrészt a TDMP területén született tételek a többváltozós Lagrange interpoláció elméletéhez is új eredményeket szolgáltatnak.

A TDMP felvetése és tárgyalása Prékopa [6, 8, 9] cikkeihez fűződik. A problémát, az  $(X_1, \dots, X_s)$  valószínűségi vektorváltozóra fogalmazzuk meg. Feltesszük,



hogy  $X_j$  értékkészlete egy ismert véges halmaz:  $Z_j = \{z_{j0}, \dots, z_{jn_j}\}$ ,  $j = 1, \dots, s$ . Vezessük be a következő jelölést:

$$p_{i_1 \dots i_s} = P(X_1 = z_{1i_1}, \dots, X_s = z_{si_s}), \quad 0 \leq i_j \leq n_j, \quad j = 1, \dots, s,$$

$$\mu_{\alpha_1 \dots \alpha_s} = \sum_{i_1=0}^{n_1} \dots \sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1} \dots z_{si_s}^{\alpha_s} p_{i_1 \dots i_s},$$

ahol  $\alpha_1, \dots, \alpha_s$  nemnegatív egész számok. A  $\mu_{\alpha_1, \dots, \alpha_s}$  számot az  $(X_1, \dots, X_s)$  valószínűségi vektorváltozó  $(\alpha_1, \dots, \alpha_s)$  rendű (hatvány)momentumának nevezzük. Az  $\alpha_1 + \dots + \alpha_s$  összeg a momentum teljes rendje.

Az alábbiakban a lineáris programozással kapcsolatos legfontosabb tények ismereteit feltételezzük. Az e tekintetben tájékozatlan olvasó figyelmébe ajánljuk Prékopa [7] cikkét, amely ezeket az ismereteket röviden összefoglalja.

Tekintsük az  $f(z)$ ,  $z \in Z$ ,  $Z = Z_1 \times \dots \times Z_s$  függvényt és vezessük be a következő jelölést:  $f_{i_1 \dots i_s} = f(z_{1i_1}, \dots, z_{si_s})$ . A többváltozós diszkrét momentum probléma egyik megfogalmazása a következő lineáris programozási feladat:

$$(2.1) \quad \min (\max) \sum_{i_1=0}^{n_1} \dots \sum_{i_s=0}^{n_s} f_{i_1 \dots i_s} p_{i_1 \dots i_s}$$

feltéve, hogy

$$\sum_{i_1=0}^{n_1} \dots \sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1} \dots z_{si_s}^{\alpha_s} p_{i_1 \dots i_s} = \mu_{\alpha_1 \dots \alpha_s}$$

$$\alpha_j \geq 0, \quad j = 1, \dots, s; \quad \alpha_1 + \dots + \alpha_s \leq m$$

$$p_{i_1 \dots i_s} \geq 0, \quad \text{minden } i_1, \dots, i_s \text{ esetén.}$$

A fenti problémát általánosíthatjuk, ha bevezetünk néhány, az  $m$  számnál magasabb rendű egyváltozós momentumra vonatkozó korlátozó feltételt is. Ennek egyik lehetséges módját fejezi ki a következő modell:

$$(2.2) \quad \min (\max) \sum_{i_1=0}^{n_1} \dots \sum_{i_s=0}^{n_s} f_{i_1 \dots i_s} p_{i_1 \dots i_s}$$

feltéve, hogy

$$\sum_{i_1=0}^{n_1} \dots \sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1} \dots z_{si_s}^{\alpha_s} p_{i_1 \dots i_s} = \mu_{\alpha_1 \dots \alpha_s}$$

$$\alpha_j \geq 0, \quad j = 1, \dots, s; \quad \alpha_1 + \dots + \alpha_s \leq m \quad \text{és}$$

$$\alpha_j = 0, \quad j = 1, \dots, k-1, k+1, \dots, s, \quad m \leq \alpha_k \leq m_k, \quad k = 1, \dots, s;$$

$$p_{i_1 \dots i_s} \geq 0, \quad \text{minden } i_1, \dots, i_s \text{ esetén.}$$

Az (2.1) és a (2.2) problémákban a  $p_{i_1 \dots i_s}$  szimbólumok a változók, míg a többi érték adott.

A fenti problémák korlátokat szolgáltatnak az

$$(2.3) \quad E[f(X_1, \dots, X_s)]$$

értékre, a megfelelő momentumok ismeretében. A (2.3) várható érték speciális esetei, alkalmas  $f$  függvények esetén, a

$$(2.4) \quad P(X_1 \geq r_1, \dots, X_s \geq r_s)$$

és a

$$(2.5) \quad P(X_1 = r_1, \dots, X_s = r_s),$$

valószínűségeket, ahol  $(r_1, \dots, r_s) \in Z$ .

A (2.1) és a (2.2) feladatok egyszerűbb alakban is felírhatók, mátrixok tenzorszorzatának segítségével. Az  $m_1 \times n_1$  méretű  $B = (b_{ij})$  mátrix és az  $m_2 \times n_2$  méretű  $C = (c_{ij})$  mátrix tenzorszorzata a  $B \otimes C$ ,  $m_1 m_2 \times n_1 n_2$  méretű mátrix, ahol  $B \otimes C = (c_{ij} B)$ . Ismert (lásd pl. Horn és Johnson [1]), hogy a tenzorszorzat asszociatív, de nem kommutatív. Vezessük be a következő jelöléseket:

$$A_j = \begin{pmatrix} 1 & 1 & \dots & 1 \\ z_{j0} & z_{j1} & \dots & z_{jn_j} \\ \vdots & & \ddots & \\ z_{j0}^{m_j} & z_{j1}^{m_j} & \dots & z_{jn_j}^{m_j} \end{pmatrix},$$

$$A = A_1 \otimes \dots \otimes A_s,$$

$$\mathbf{b} = E[(1, X_1, \dots, X_1^{m_1}) \otimes \dots \otimes (1, X_s, \dots, X_s^{m_s})]^T$$

$$= (\mu_{00\dots 0}, \mu_{10\dots 0}, \dots, \mu_{m_1 0\dots 0}, \mu_{010\dots 0}, \mu_{11\dots 0}, \dots)^T$$

$$\mathbf{p} = (p_{i_1 \dots i_s}, 0 \leq i_1 \leq n_1, \dots, 0 \leq i_s \leq n_s)^T$$

$$\mathbf{f} = (f_{i_1 \dots i_s}, 0 \leq i_1 \leq n_1, \dots, 0 \leq i_s \leq n_s)^T,$$

ahol  $\mathbf{p}$  és  $\mathbf{f}$  komponenseinek a sorrendje megegyezik az  $A$  mátrix megfelelő oszlopainak sorrendjével. Az  $A$  megfelelő sorainak, illetve a  $\mathbf{b}$  megfelelő komponenseinek

a kiválasztásával, a fenti problémák tömör alakban is felírhatók. Az (2.1) feladat tömör formája:

$$(2.6) \quad \min(\max) \quad \mathbf{f}^T \mathbf{p}$$

feltéve, hogy

$$\tilde{A}\mathbf{p} = \tilde{\mathbf{b}}$$

$$\mathbf{p} \geq \mathbf{0},$$

a (2.2) feladaté pedig:

$$(2.7) \quad \min(\max) \quad \mathbf{f}^T \mathbf{p}$$

feltéve, hogy

$$\hat{A}\mathbf{p} = \hat{\mathbf{b}}$$

$$\mathbf{p} \geq \mathbf{0}.$$

Az  $A$  mátrix mérete  $[(m_1 + 1) \cdots (m_s + 1)] \times [(n_1 + 1) \cdots (n_s + 1)]$ , míg az  $\tilde{A}$  mátrixé  $N \times [(n_1 + 1) \cdots (n_s + 1)]$ , ahol  $N = \binom{s+m}{m}$ . Az  $\hat{A}$  mátrix mérete  $N' \times [(n_1 + 1) \cdots (n_s + 1)]$ , ahol  $N' = N + \sum_{j=1}^s (m_j - m)$ . Az  $\tilde{A}$  mátrix teljes rangú, ha  $m \leq n_j$ ,  $j = 1, \dots, s$ ; az  $\hat{A}$  mátrix teljes rangú, ha  $m_j \leq n_j$ ,  $j = 1, \dots, s$ .

Jelölje  $V_{\min}$  ( $V_{\max}$ ) a (2.1) vagy (2.2) probléma minimumát (maximumát). Jelölje továbbá  $B_1$  ( $B_2$ ) a minimum (maximum) probléma egy duál megengedett bázisát (olyan bázist, melyre az optimalitási feltételek teljesülnek). Ezek után, a lineáris programozás elméletét felhasználva, felírhatjuk a következő egyenlőtlenségeket:

$$(2.8) \quad \mathbf{f}_{B_1}^T \mathbf{p}_{B_1} \leq V_{\min} \leq E[f(X_1, \dots, X_s)] \leq V_{\max} \leq \mathbf{f}_{B_2}^T \mathbf{p}_{B_2}.$$

Ha  $B_1$  ( $B_2$ ) primál megengedett is, tehát optimális, bázis a minimum (maximum) problémában, akkor az első (utolsó) egyenlőtlenség egyenlőséggel teljesül. Ekkor azt mondjuk, hogy  $V_{\min}$  és  $V_{\max}$  éles alsó és felső korlátok az  $f(X_1, \dots, X_s)$  függvény várható értékére.

Végül, hogy kapcsolatot teremthessünk a többváltozós Lagrange interpoláció és a (2.6), (2.7) problémák között, bevezetjük a következő terminológiát.

Legyen az  $U = \{\mathbf{u}_1, \dots, \mathbf{u}_M\}$  az  $\mathbb{R}^s$  tér egy halmaza, és  $H = \{(\alpha_1, \dots, \alpha_s)\}$  az  $(\alpha_1, \dots, \alpha_s)$  nemnegatív egész komponensekből alkotott  $s$  dimenziós vektorok egy véges halmaza.

Azt mondjuk, hogy  $U$  megenged egy  $H$ -típusú Lagrange interpolációt, ha bármilyen valós  $f(z)$ ,  $z \in U$  függvény esetén létezik olyan  $p(z)$  polinom, mely a következő alakban írható,

$$(2.9) \quad p(z) = \sum_{(\alpha_1, \dots, \alpha_s) \in H} c(\alpha_1, \dots, \alpha_s) z_1^{\alpha_1} \cdots z_s^{\alpha_s},$$

ahol  $c(\alpha_1, \dots, \alpha_s)$  valós számok, és

$$(2.10) \quad p(u_i) = f(u_i), \quad i = 1, \dots, M.$$

Definiáljuk a  $\tilde{b}(z_1, \dots, z_s)$  ( $\widehat{b}(z_1, \dots, z_s)$ ) vektort a  $\tilde{b}$  ( $\widehat{b}$ ) vektorhoz hasonlóan, azzal a különbséggel, hogy elhagyjuk a várható értéket és az  $X_j$  valószínűségi változó helyére a  $z_j$  determinisztikus változót írjuk,  $j = 1, \dots, s$ .

A (2.6) ((2.7)) probléma esetén definiáljuk a  $H$ ,  $I$  és  $U$  halmazokat a következő módon:

$$(2.11) \quad H = \{(\alpha_1, \dots, \alpha_s) \mid 0 \leq \alpha_j, \alpha_j \text{ egész}, \alpha_1 + \dots + \alpha_s \leq m, j = 1, \dots, s\}$$

$$(H = \{(\alpha_1, \dots, \alpha_s) \mid 0 \leq \alpha_j, \alpha_j \text{ egész}, \alpha_1 + \dots + \alpha_s \leq m, j = 1, \dots, s;$$

$$\text{vagy } \alpha_j = 0, j = 1, \dots, k-1, k+1, \dots, s, \quad m \leq \alpha_k \leq m_k,$$

$$k = 1, \dots, s\}),$$

$$I = \{(i_1, \dots, i_s) \mid \tilde{a}_{i_1 \dots i_s} \in \tilde{B}\}$$

$$(I = \{(i_1, \dots, i_s) \mid \widehat{a}_{i_1 \dots i_s} \in \widehat{B}\}),$$

$$U = \{(z_{1i_1}, \dots, z_{si_s}) \mid (i_1, \dots, i_s) \in I\}.$$

Ekkor az

$$(2.12) \quad L_I(z_1, \dots, z_s) = \mathbf{f}_B^T \tilde{B}^{-1} \tilde{b}(z_1, \dots, z_s)$$

$$(L_I(z_1, \dots, z_s) = \mathbf{f}_B^T \widehat{B}^{-1} \widehat{b}(z_1, \dots, z_s))$$

polinom megegyezik az  $U$  halmazhoz tartozó  $H$ -típusú, egyértelműen adott Lagrange polinommal.

A  $\tilde{B}$  illetve  $\widehat{B}$  bázis duál megengedettsége a minimum (maximum) problémára a következőket jelenti:

$$(2.13) \quad f(z_1, \dots, z_s) \geq L_I(z_1, \dots, z_s), \quad \text{minden } (z_1, \dots, z_s) \in Z \quad \text{esetén}$$

$$(f(z_1, \dots, z_s) \leq L_I(z_1, \dots, z_s), \quad \text{minden } (z_1, \dots, z_s) \in Z \quad \text{esetén}),$$

ahol  $(z_1, \dots, z_s) \in U$  esetén egyenlőség áll.

A (2.13) relációt a minimum (maximum) probléma optimalitási feltételének nevezzük. Ha a (2.13) relációkban  $(z_1, \dots, z_s)$  helyére az  $(X_1, \dots, X_s)$  valószínűségi vektorváltozót helyettesítjük és vesszük a kifejezés várható értékét, akkor korlátot kapunk az  $E[f(X_1, \dots, X_s)]$  várható értékre. Ha a bázis egyben primál megengedett is, akkor a kapott korlát éles.

A fentiek szerint a TDMP problémák alkalmasak arra, hogy a célfüggvény valódi értékére korlátokat adjunk duál megengedett bázisok segítségével. Másrészt, egy bázis duál megengedettsége, (2.13) szerint, egyenértékű azzal, hogy létezik a bázisvektoroknak megfelelő tartópontokhoz olyan, a momentumok rendje által meghatározott típusú Lagrange interpolációs polinom, mely a tartó többi pontján nem nagyobb (nem kisebb) mint a függvényérték. Tehát azt mondhatjuk, hogy a korlátok konstrukciójára vonatkozó fenti technika lényegében abban áll, hogy bizonyos pontthalmazokra olyan megfelelő típusú interpolációs polinomot konstruáljunk, melyre az  $R_I = f - L_I$  maradéktag mindig nemnegatív (nempozitív).

### 3. Egy tétel többváltozós Lagrange interpolációra

Tekintsük a következő indexhalmazt:

$$(3.1) \quad I = I_0 \cup \left( \bigcup_{j=1}^s I_j \right),$$

ahol

$$(3.2) \quad I_0 = \{(i_1, \dots, i_s) \mid 0 \leq i_j \leq m-1, \text{ egészek, } j = 1, \dots, s, i_1 + \dots + i_s \leq m\}$$

és

$$(3.3) \quad I_j = \{(i_1, \dots, i_s) \mid i_j \in K_j, i_l = 0 \text{ } l \neq j\}$$

$$K_j = \{k_j^{(1)}, \dots, k_j^{(|K_j|)}\} \subset \{m, m+1, \dots, n_j\}, \quad j = 1, \dots, s.$$

Vezessük be az alábbi jelöléseket:

$$Z_{ji} = \{z_{j0}, \dots, z_{ji}\},$$

$$Z'_{ji} = \{z_{j0}, \dots, z_{ji}, z_j\},$$

$$i = 0, \dots, n_j, \quad j = 1, \dots, s;$$

$$K_{ji} = \{k_j^{(1)}, \dots, k_j^{(i)}\},$$

$$Z_{jK_j} = \{z_{jk_j^{(1)}}, \dots, z_{jk_j^{(i)}}\},$$

$$i = 1, \dots, |K_j|, \quad j = 1, \dots, s,$$

$$Z_{jK_j} = Z_{jK_{|K_j|}}, \quad j = 1, \dots, s.$$

A  $Z_I = \{(z_{1i_1}, \dots, z_{si_s}) \mid (i_1, \dots, i_s) \in I\}$  pontokhoz tartozó Lagrange polinomot az alábbi Newton-féle alakkal adjuk meg:

$$(3.4) \quad L_I(z_1, \dots, z_s) =$$

$$= \sum_{\substack{i_1 + \dots + i_s \leq m \\ 0 \leq i_j \leq m-1, j=1, \dots, s}} [Z_{1i_1}; \dots; Z_{si_s}; f] \prod_{j=1}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}) +$$

$$+ \sum_{j=1}^s \sum_{i=1}^{|K_j|} [Z_{10}; \dots; Z_{(j-1)0}; Z_{j(m-1)} \cup Z_{jK_j}; Z_{(j+1)0}; \dots; Z_{s0}; f] \times$$

$$\times \prod_{k \in \{0, \dots, m-1\} \cup K_{j(i-1)}} (z_j - z_{jk}),$$

ahol, definíció szerint  $\prod_{k=0}^{i_j-1} (z_j - z_{jk}) = 1$ , ha  $i_j = 0$ , és  $K_{j0} = \emptyset$ .

A (3.4) kifejezésben az  $f$  függvény bármilyen függvényt jelölhet, melynek a  $Z$  halmaz elemeihez tartozó értékei megegyeznek az eredeti diszkrét függvény értékeivel.

A maradéktagot a következő módon adjuk meg:

$$(3.5) \quad R_I(z_1, \dots, z_s) = R_{1I}(z_1, \dots, z_s) + R_{2I}(z_1, \dots, z_s),$$

ahol

$$(3.6) \quad R_{1I}(z_1, \dots, z_s) =$$

$$= \sum_{j=1}^s [z_{10}; \dots; z_{(j-1)0}; Z_{j(m-1)} \cup Z_{jK_j} \cup \{z_j\}; z_{(j+1)0}; \dots; z_{s0}; f] \times$$

$$\times \prod_{k \in \{0, \dots, m-1\} \cup K_j} (z_j - z_{jk})$$

és

$$(3.7) \quad R_{2I}(z_1, \dots, z_s) =$$

$$\begin{aligned}
 &= \sum_{h=1}^s \sum_{\substack{i_h + \dots + i_s = m \\ 0 \leq i_j \leq m-1, j=h, \dots, s}} \left[ z_1; \dots; z_{h-1}; Z'_{hi_h}; Z_{(h+1)i_{h+1}}; \dots; Z_{si_s}; f \right] \times \\
 &\quad \times \prod_{l=0}^{i_h} (z_h - z_{hl}) \prod_{h+1}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}) + \\
 &+ \sum_{j=h+1}^s \left[ z_1; \dots; z_{h-1}; Z'_{h0}; Z_{(h+1)0}; \dots; Z_{(j-1)0}; Z'_{j(m-1)}; Z_{(j+1)0}; \dots; Z_{s0} \right] \times \\
 &\quad \times (z_h - z_{h0}) \prod_{k=0}^{m-1} (z_j - z_{jk}).
 \end{aligned}$$

A következő tétel általánosítása mind az egyváltozós (1.8) formulának, mind a Prékopa [8, Section 4] cikkében található többváltozós formulának.

3.1. TÉTEL. *Tetszőleges, az  $f$  függvény értelmezési tartományába eső,  $z = (z_1, \dots, z_s)$  pontra igaz a következő egyenlőség:*

$$(3.8) \quad L_I(z_1, \dots, z_s) + R_I(z_1, \dots, z_s) = f(z_1, \dots, z_s).$$

*Bizonyítás.* Az általánosság megszorítása nélkül feltehetjük, hogy  $K_j = \{m, m+1, \dots, m_j\}$ , ahol  $m_j \geq m$ ,  $j = 1, \dots, s$ . Valóban, ha tekintjük a  $\overline{Z}_j = Z_{j(m-1)} \cup Z_{jK_j}$ ,  $j = 1, \dots, s$ , halmazt, majd erre bebizonyítjuk az állítást, akkor lényegében az eredeti általános esetre vonatkozó állítást láttuk be.

A  $K_j$ ,  $j = 1, \dots, s$ , halmazokra tett feltétel szerint az  $L_I(z_1, \dots, z_s)$  és  $R_{1I}(z_1, \dots, z_s)$  függvények a következő formát öltik:

$$\begin{aligned}
 (3.9) \quad L_I(z_1, \dots, z_s) &= \\
 &= \sum_{\substack{i_1 + \dots + i_s \leq m \\ 0 \leq i_j \leq m-1, j=1, \dots, s}} [Z_{1i_1}; \dots; Z_{si_s}; f] \prod_{j=1}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}) + \\
 &+ \sum_{j=1}^s \sum_{i_j=m}^{m_j} [Z_{10}; \dots; Z_{(j-1)0}; Z_{ji_j}; Z_{(j+1)0}; \dots; Z_{s0}; f] \prod_{k=0}^{i_j-1} (z_j - z_{jk})
 \end{aligned}$$

és

$$\begin{aligned}
 (3.10) \quad R_{1I}(z_1, \dots, z_s) &= \sum_{j=1}^s [Z_{10}; \dots; Z_{(j-1)0}; Z'_{jm_j}; Z_{(j+1)0}; \dots; Z_{s0}; f] \times \\
 &\quad \times \prod_{k=0}^{m_j} (z_j - z_{jk}).
 \end{aligned}$$

Az  $R_{2I}(z_1, \dots, z_s)$  függvényre adott képlet változatlan marad. Először a következő állítást látjuk be:

3.1. LEMMA. *Érvényes az alábbi egyenlőség:*

$$(3.11) \quad \begin{aligned} & L_I(z_1, \dots, z_s) + R_{1I}(z_1, \dots, z_s) = \\ &= \sum_{\substack{i_1 + \dots + i_s \leq m \\ 0 \leq i_j \leq m-1, j=1, \dots, s}} [Z_{1i_1}; \dots; Z_{si_s}; f] \prod_{j=1}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}) + \\ &+ \sum_{j=1}^s [Z_{10}; \dots; Z_{(j-1)0}; Z'_{j(m-1)}; Z_{(j+1)0}; \dots; Z_{s0}; f] \prod_{k=0}^{m-1} (z_j - z_{jk}). \end{aligned}$$

*Bizonyítás.* Tekintsük a  $z_j$  változó alábbi függvényét:

$$[Z_{10}; \dots; Z_{(j-1)0}; Z'_{j(m-1)}; Z_{(j+1)0}; \dots; Z_{s0}; f].$$

Az ezzel a függvénnyel vett,  $z_{jm}, \dots, z_{jm_j}$  pontokhoz tartozó Lagrange polinom így írható:

$$\sum_{i_j=m}^{m_j} [Z_{10}; \dots; Z_{(j-1)0}; Z_{ji_j}; Z_{(j+1)0}; \dots; Z_{s0}; f] \prod_{k=m}^{i_j-1} (z_j - z_{jk}).$$

Ezek után, az (1.8) képletet használva, felírhatjuk, hogy

$$(3.12) \quad \begin{aligned} & [Z_{10}; \dots; Z_{(j-1)0}; Z'_{j(m-1)}; Z_{(j+1)0}; \dots; Z_{s0}; f] = \\ &= \sum_{i_j=m}^{m_j} [Z_{10}; \dots; Z_{(j-1)0}; Z_{ji_j}; Z_{(j+1)0}; \dots; Z_{s0}; f] \prod_{k=m}^{i_j-1} (z_j - z_{jk}) + \\ &+ [Z_{10}; \dots; Z_{(j-1)0}; Z'_{jm_j}; Z_{(j+1)0}; \dots; Z_{s0}; f] \prod_{k=m}^{m_j} (z_j - z_{jk}). \end{aligned}$$

Ha a (3.12) egyenlőség minden sorát megszorozzuk a  $\prod_{k=0}^{m-1} (z_j - z_{jk})$  tényezővel, majd ezeket összegezzük  $j$  szerint, akkor a következőt kapjuk:

$$(3.13) \quad \begin{aligned} & \sum_{j=1}^s [Z_{10}; \dots; Z_{(j-1)0}; Z'_{j(m-1)}; Z_{(j+1)0}; \dots; Z_{s0}; f] \prod_{k=0}^{m-1} (z_j - z_{jk}) = \\ &= \sum_{j=1}^s \sum_{i_j=m}^{m_j} [Z_{10}; \dots; Z_{(j-1)0}; Z_{ji_j}; Z_{(j+1)0}; \dots; Z_{s0}; f] \prod_{k=0}^{i_j-1} (z_j - z_{jk}) + \\ &+ R_{1I}(z_1, \dots, z_s). \end{aligned}$$



A (3.9) és (3.13) képletekből már adódik a lemma állítása.  $\square$

Ha (3.11) harmadik sorában külön vesszük a  $j = 1$  tagot, akkor a következő formulát kapjuk:

$$(3.14) \quad L_{1I}(z_1, \dots, z_s) + R_{1I}(z_1, \dots, z_s) = \sum_{\substack{i_1 + \dots + i_s = m \\ 0 \leq i_j \leq m-1, j=1, \dots, s}} [Z_{1i_1}; \dots; Z_{si_s}; f] \prod_{j=1}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}) +$$

$$(3.15) \quad + [Z'_{1(m-1)}; Z_{20}; \dots; Z_{s0}; f] \prod_{k=0}^{m-1} (z_1 - z_{1k}) +$$

$$(3.16) \quad + \sum_{j=2}^s [Z_{10}; \dots; Z_{(j-1)0}; Z'_{j(m-1)}; Z_{(j+1)0}; \dots; Z_{s0}; f] \times \\ \times \prod_{k=0}^{m-1} (z_j - z_{jk}).$$

Hasonlóan, ha  $R_{2I}$  képletében leválasztjuk a  $h = 1$  tagot, azt kapjuk hogy:

$$(3.17) \quad R_{2I}(z_1, \dots, z_s) = \sum_{\substack{i_1 + \dots + i_s = m \\ 0 \leq i_j \leq m-1, j=1, \dots, s}} [Z'_{1i_1}; Z_{2i_2}; \dots; Z_{si_s}; f] \prod_{l=0}^{i_1} (z_1 - z_{1l}) \times \\ \times \prod_{j=2}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}) +$$

$$(3.18) \quad + \sum_{j=2}^s [Z'_{10}; Z_{20}; \dots; Z_{(j-1)0}; Z'_{j(m-1)}; Z_{(j+1)0}; \dots; Z_{s0}] (z_1 - z_{10}) \times \\ \times \prod_{k=0}^{m-1} (z_j - z_{jk}) +$$

$$(3.19) \quad + \sum_{h=2}^s \left( \sum_{\substack{i_h + \dots + i_s = m \\ 0 \leq i_j \leq m-1, j=h, \dots, s}} [z_1; \dots; z_{h-1}; Z'_{hi_h}; Z_{(h+1)i_{h+1}}; \dots; Z_{si_s}; f] \times \right.$$

$$\begin{aligned}
& \times \prod_{l=0}^{i_h} (z_h - z_{hl}) \prod_{h+1}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}) + \\
& + \sum_{j=h+1}^s \left[ z_1; \dots; z_{h-1}; Z'_{h0}; Z_{(h+1)0}; \dots; Z_{(j-1)0}; Z'_{j(m-1)}; Z_{(j+1)0}; \dots; Z_{s0} \right] \times \\
& \times (z_h - z_{h0}) \prod_{k=0}^{m-1} (z_j - z_{jk}).
\end{aligned}$$

A következő lépésben meghatározzuk a (3.14), (3.15) és (3.17) tagok összegét. Először felírjuk a (3.14) relációt az alábbi formában:

$$\begin{aligned}
(3.20) \quad & \sum_{\substack{0 < i_2 + \dots + i_s \leq m \\ 0 \leq i_j \leq m-1, \ j=2, \dots, s}} \left( \sum_{i_1=0}^{m-i_2-\dots-i_s} [Z_{1i_1}; Z_{2i_2}; \dots; Z_{si_s}; f] \prod_{l=0}^{i_1-1} (z_1 - z_{1l}) \right) \times \\
& \times \prod_{j=2}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}) + \\
(3.21) \quad & + \sum_{i_1=0}^{m-1} [Z_{1i_1}; Z_{20}; \dots; Z_{s0}; f] \prod_{l=0}^{i_1-1} (z_1 - z_{1l}).
\end{aligned}$$

Ezt követően a (3.17) kifejezést az alábbi alakba írjuk:

$$\begin{aligned}
(3.22) \quad & \sum_{\substack{0 < i_2 + \dots + i_s \leq m \\ 0 \leq i_j \leq m-1, \ j=2, \dots, s}} \left( [Z'_{1(m-i_1-\dots-i_s)}; Z_{2i_2}; \dots; Z_{si_s}; f] \prod_{l=0}^{i_1} (z_1 - z_{1l}) \right) \times \\
& \times \prod_{j=2}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}).
\end{aligned}$$

Célunk az, hogy összeadjuk a (3.20), (3.21), (3.22) és (3.15) formulákat. Egyrészt (3.21) és (3.15) összege (felhasználva az (1.8) képletet) a következő:

$$(3.23) \quad [z_1; Z_{20}; \dots; Z_{s0}; f].$$

Másrészt, (3.20) és (3.22) összege (először a zárójelben lévő tagokat adjuk össze) a következő:

$$(3.24) \quad \sum_{\substack{0 < i_2 + \dots + i_s \leq m \\ 0 \leq i_j \leq m-1, \ j=2, \dots, s}} [z_1; Z_{2i_2}; \dots; Z_{si_s}; f] \prod_{j=2}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}).$$

A (3.23) és (3.24) összege, és egyben a bizonyítás ezen lépésének eredménye:

$$(3.25) \quad \sum_{\substack{i_2+\dots+i_s \leq m \\ 0 \leq i_j \leq m-1, \ j=2,\dots,s}} [z_1; Z_{2i_2}; \dots; Z_{si_s}; f] \prod_{j=2}^s \prod_{k=0}^{i_j-1} (z_j - z_{jk}).$$

A következő lépés az, hogy kiszámoljuk a még figyelembe nem vett (3.16) és (3.18) tagok összegét. Tekintve a  $j$  indexhez tartozó tagokat a (3.16) és (3.18) sorokban, a  $\prod_{k=0}^{m-1} (z_j - z_{jk})$  tényező nélkül, a két tag összege (felhasználva az (1.8) képletet a  $z_1$  változóra) az alábbi:

$$(3.26) \quad [z_1; Z_{20}; \dots; Z_{(j-1)0}; Z'_{j(m-1)}; Z_{(j+1)0}; \dots; Z_{s0}; f].$$

Eszerint a (3.16) és (3.18) képletek összege az alábbival egyenlő:

$$(3.27) \quad \sum_{j=2}^s [z_1; Z_{20}; \dots; Z_{(j-1)0}; Z'_{j(m-1)}; Z_{(j+1)0}; \dots; Z_{s0}; f] \prod_{k=0}^{m-1} (z_j - z_{jk}).$$

Idáig azt az eredményt kaptuk, hogy az  $L_I(z_1, \dots, z_s) + R_{1I}(z_1, \dots, z_s) + R_{2I}(z_1, \dots, z_s)$  összeg megegyezik a (3.25), (3.27) és (3.19) formulák összegével. Defináljuk a  $J$  indexhalmazt az  $I$  halmazhoz hasonlóan, de csak  $i_2, \dots, i_s$ -re vonatkoztatva, és tekintsük az  $s-1$  változós  $f(z_1, z_2, \dots, z_s)$  függvényt, ahol  $z_1 \in Z$  kötött. Ekkor a (3.25) és a (3.27) képlet összege egyenlő a  $L_J + R_{1J}$  összeggel, míg (3.19) egyenlő az  $R_{2J}$  taggal. Ha feltesszük, hogy (3.8) igaz bármely  $s-1$  változós függvényre, akkor, a fenti érvelés szerint, igaz minden  $s$  változós függvényre is.  $\square$

#### 4. Diszkrét, magasabb rendű konvex függvények féloldalas approximációja

Tekintsünk egy  $f(z)$ ,  $z \in Z$  függvényt, mellyel kapcsolatban megkövetelünk bizonyos konvexitási feltételeket. A 3.1. Tételre támaszkodva megadjuk a  $Z$  halmaz egy olyan részhalmazát, melyre az előző fejezet Lagrange polinomja a függvény egy alsó illetve felső közelítését szolgáltatja. Egyidejűleg a megfelelő TDMP modellek egy duál megengedett bázisát is megkapjuk, figyelembe véve a 2. fejezetben elmondottakat.

Jelölje  $H$  a (2.11) zárójelében definiált halmazt. A 3. szakaszban bevezetett  $K_j \subset \{m, m+1, \dots, n_j\}$  halmazokra vonatkozólag az alábbi struktúrák valamelyikének az érvényességét fogjuk megkívánni:

$$\begin{aligned}
 (4.1) \quad & \begin{array}{ll} |K_j| & \text{páros} \\ \min & u^{(j)}, u^{(j)} + 1, \dots, v^{(j)}, v^{(j)} + 1 \\ \max & m, u^{(j)}, u^{(j)} + 1, \dots, v^{(j)}, v^{(j)} + 1, n_j \end{array} \\
 & \begin{array}{ll} |K_j| & \text{páratlan} \\ \min & m, u^{(j)}, u^{(j)} + 1, \dots, v^{(j)}, v^{(j)} + 1 \\ \max & u^{(j)}, u^{(j)} + 1, \dots, v^{(j)}, v^{(j)} + 1, n_j. \end{array}
 \end{aligned}$$

Az eddigiekben nem tételeztük fel, hogy a  $Z_1, \dots, Z_s$  halmazok elemei rendezettek. A most következő tételekben azonban elemeiket növekvő vagy csökkenő sorrendben rendezzük el.

4.1. TÉTEL. Legyen  $z_{j0} < z_{j1} < \dots < z_{jn_j}$ ,  $j = 1, \dots, s$ . Tegyük fel, hogy az  $f(z)$ ,  $z \in Z$  függvény  $m + 1$  rendű osztott differenciái nemnegatívak, továbbá, hogy a  $z_j$  változója szerinti  $m + |K_j|$  rendű osztott differenciái is nemnegatívak, ahol  $K_j$  (4.1) valamelyik min struktúráját követi,  $j = 1, \dots, s$ .

Ekkor igaz az, hogy a (3.4) képlettel definiált  $L_I(z_1, \dots, z_s)$  megegyezik a  $Z_I$  halmazhoz tartozó,  $H$ -típusú, egyértelműen adott Lagrange polinommal, és teljesíti az alábbi egyenlőtlenséget:

$$(4.2) \quad f(z_1, \dots, z_s) \geq L_I(z_1, \dots, z_s), \quad (z_1, \dots, z_s) \in Z.$$

Eszerint, a (2.7) minimum probléma  $\hat{A}$  mátrixának  $I$  indexhalmazhoz tartozó oszlopaiból álló  $\hat{B}$  bázis duál megengedett, és

$$(4.3) \quad E[f(X_1, \dots, X_s)] \geq E[L_I(X_1, \dots, X_s)].$$

Ha  $\hat{B}$  egyben primál megengedett is, akkor a (4.3) képlettel adott korlát éles.

Ha a fent említett osztott differenciák mindegyike nempozitív, akkor (4.2) és (4.3) fordított relációjelekkel állnak.

*Bizonyítás.* A (3.9)  $H$ -típusú Lagrange polinom egyértelműségét, illetve azt, hogy  $\hat{B}$  a (2.7) LP feladat bázisa, a következőképp láthatjuk be. Abból, hogy  $(z = (z_1, \dots, z_s))$   $L_I(z) = f(z)$ ,  $z \in Z_I$  esetén, következik, hogy  $f_{\hat{B}}$  előáll  $\hat{B}$  sorainak lineáris kombinációiként. Mivel ez bármilyen  $f$  függvényre, és így bármilyen  $f_{\hat{B}}$  vektorra igaz, ezért a  $\hat{B}$  mátrix nem szinguláris. Ebből már következik a Lagrange polinom egyértelműsége.

Azt, hogy a (4.2) egyenlőtlenség ekvivalens a  $\hat{B}$  bázis (2.7) feladatbeli duál megengedettséggel, a 2. fejezetben már levezettük.

A (4.2) reláció bizonyításához használjuk a (3.8) egyenletet. Mivel  $R_I(z_1, \dots, z_s) = R_{1I}(z_1, \dots, z_s) + R_{2I}(z_1, \dots, z_s)$ , ezért elég, ha belátjuk, hogy  $R_{1I}(z_1, \dots, z_s) \geq 0$  és  $R_{2I}(z_1, \dots, z_s) \geq 0$  teljesül  $(z_1, \dots, z_s) \in Z$  esetén.

Az  $R_{1I}(z_1, \dots, z_s)$  függvény (3.6) képlettel adott alakjából és  $K_j$  speciális struktúrájából következik, hogy

$$(4.4) \quad \prod_{k \in \{0, \dots, m-1\} \cup K_j} (z_j - z_{jk}) > 0, \quad \text{ha } j \notin \{0, \dots, m-1\} \cup K_j.$$

Ha pedig  $j \in \{0, \dots, m-1\} \cup K_j$ , akkor a fenti szorzat eltűnik. Mivel a feltétel szerint az  $f$  függvény  $z_j$ ,  $j = 1, \dots, s$  változó szerinti  $m + |K_j|$  rendű osztott differenciái nemnegatívak, ezért  $R_{1I}(z_1, \dots, z_s) \geq 0$  minden  $(z_1, \dots, z_s) \in Z$  esetén.

Az  $R_{2I}(z_1, \dots, z_s)$  függvény (3.7) képletével adott összeg tagjaiban mind az osztott differenciák, melyek teljes rendje  $m + 1$ , mind pedig azok szorzói nemnegatívak bármely  $(z_1, \dots, z_s) \in Z$  esetén. Ebből következik, hogy  $R_{2I}(z_1, \dots, z_s) \geq 0$ ,  $(z_1, \dots, z_s) \in Z$ , ami egyenértékű a (4.2) relációval. Ebből viszont (4.3) reláció érvényessége közvetlenül adódik.

Végül, ha  $\widehat{B}$  egyszerre primál és duál megengedett bázisa a (2.7) problémának, akkor ez optimális bázis, és a célfüggvény értéke az alábbi:

$$\begin{aligned} \min E[f(z_1, \dots, z_s)] &= \\ &= f_B^T \mathbf{p}_b = f_B^T \widehat{B}^{-1} \widehat{\mathbf{b}} = \\ &= f_B^T \widehat{B}^{-1} E[\widehat{\mathbf{b}}(X_1, \dots, X_s)] = \\ &= E[f_B^T \widehat{B}^{-1} \widehat{\mathbf{b}}(X_1, \dots, X_s)] = \\ &= E[L_I(X_1, \dots, X_s)]. \end{aligned} \quad \square$$

A következő tétel a fenti feltételek módosítása mellett szolgáltatja az  $f(z_1, \dots, z_s)$ ,  $(z_1, \dots, z_s) \in Z$  függvénynek alsó illetve felső approximációját, egyben korlátokat is ad annak várható értékére.

**4.2. TÉTEL.** Legyen  $z_{j0} > z_{j1} > \dots > z_{jn_j}$ ,  $j = 1, \dots, s$ . Tegyük fel, hogy az  $f(\mathbf{z})$ ,  $\mathbf{z} \in Z$  függvény  $m + 1$  rendű osztott differenciái nemnegatívak, továbbá, hogy a  $z_j$  változója szerinti  $m + |K_j|$  rendű osztott differenciái ugyancsak nemnegatívak, ahol  $K_j$  (4.1) valamelyik, az alábbiakban jelzett, struktúráját követi,  $j = 1, \dots, s$ . Ekkor igazak a következők:

(i) Ha  $m + 1$  páros,  $|K_j|$  páros és  $K_j$  (4.1) megfelelő max struktúráját követi, vagy ha  $m + 1$  páros,  $|K_j|$  páratlan és  $K_j$  (4.1) megfelelő min struktúráját követi, akkor a (3.4) képlettel definiált  $L_I(z_1, \dots, z_s)$  megegyezik a  $Z_I$  halmazhoz tartozó,  $H$ -típusú, egyértelműen adott Lagrange polinommal, és teljesíti az alábbi egyenlőtlenséget:

$$(4.5) \quad f(z_1, \dots, z_s) \geq L_I(z_1, \dots, z_s), \quad (z_1, \dots, z_s) \in Z.$$

Eszerint a (2.7) minimum probléma  $\hat{A}$  mátrixának  $I$  indexhalmazhoz tartozó oszlopaiból álló  $\hat{B}$  bázis duál megengedett, és

$$(4.6) \quad E[f(X_1, \dots, X_s)] \geq E[L_I(X_1, \dots, X_s)].$$

Ha  $\hat{B}$  egyben primál megengedett is, akkor a (4.6) képlettel adott korlát éles.

(ii) Ha  $m + 1$  páratlan,  $|K_j|$  páros és  $K_j$  (4.1) megfelelő max struktúráját követi, vagy ha  $m + 1$  páratlan,  $|K_j|$  páratlan és  $K_j$  (4.1) megfelelő min struktúráját követi, akkor a (3.4) képlettel definiált  $L_I(z_1, \dots, z_s)$  megegyezik a  $Z_I$  halmazhoz tartozó,  $H$ -típusú, egyértelműen adott Lagrange polinommal, és teljesíti az alábbi egyenlőtlenséget:

$$(4.7) \quad f(z_1, \dots, z_s) \leq L_I(z_1, \dots, z_s), \quad (z_1, \dots, z_s) \in Z.$$

Eszerint a (2.7) maximum probléma  $\hat{A}$  mátrixának  $I$  indexhalmazhoz tartozó oszlopaiból álló  $\hat{B}$  bázis duál megengedett, és

$$(4.8) \quad E[f(X_1, \dots, X_s)] \leq E[L_I(X_1, \dots, X_s)].$$

Ha  $\hat{B}$  egyben primál megengedett is, akkor a (4.8) képlettel adott korlát éles.

*Bizonyítás.* Az állítások az előző tétel bizonyításában használt módszerek segítségével már könnyen beláthatók.  $\square$

Mint az előzőekben is utaltunk rá, a függvény várható értékének éles alsó (felső) korlátait a megfelelő TDMP minimum (maximum) probléma optimális célfüggvény értéke szolgáltatja. A TDMP feladatot megoldhatjuk duál szimplex algoritmussal. Ebben az esetben a fenti állításokban szereplő duál megengedett bázisok, azontúl, hogy korlátokat szolgáltatnak a függvény várható értékére, a duál algoritmus kezdeti bázisaiként is nagy segítséget nyújthatnak. Egy kezdeti duál megengedett bázis ismerete két lényeges előnnyel is jár: egyrészt a duál algoritmus első fázisának elhagyásával megtakarítjuk a futási idő körülbelül felét, másrészt a kevesebb művelet miatt a numerikus pontosság is jobb lesz.

Befejezésül, nézzünk egy-egy példát alulról illetve felülről approximáló interpolációs polinomokra kétdimenziós esetben.

**4.1. Példa.** Tegyük fel, hogy az  $f(z_1, z_2)$ ,  $z_1 \in Z_1$ ,  $z_2 \in Z_2$  függvény 3 rendű osztott differenciái nemnegatívak, továbbá, hogy a  $z_j$  változója szerinti 5 rendű osztott differenciái ugyancsak nemnegatívak,  $j = 1, 2$ . Ebben az esetben ( $m = 2$ ,  $|K_1| = |K_2| = 3$ ) a 4.1. Tétel a következő alulról közelítő, Lagrange polinomot adja:

$$(4.9) \quad L_I(z_1, z_2) = \\ = [z_{10}; z_{20}; f] + [z_{10}, z_{11}; z_{20}; f](z_1 - z_{10}) +$$

$$\begin{aligned}
& + [z_{10}; z_{20}, z_{21}; f](z_2 - z_{20}) + \\
& + [z_{10}, z_{11}; z_{20}, z_{21}; f](z_1 - z_{10})(z_2 - z_{20}) + \\
& + [z_{10}, z_{11}, z_{12}; z_{20}; f](z_1 - z_{10})(z_1 - z_{11}) + \\
& + [z_{10}, z_{11}, z_{12}, z_{1i}; z_{20}; f](z_1 - z_{10})(z_1 - z_{11})(z_1 - z_{12}) + \\
& + [z_{10}, z_{11}, z_{12}, z_{1i}, z_{1(i+1)}; z_{20}; f](z_1 - z_{10})(z_1 - z_{11})(z_1 - z_{12})(z_1 - z_{1i}) + \\
& + [z_{10}; z_{20}, z_{21}, z_{22}; f](z_2 - z_{20})(z_2 - z_{21}) + \\
& + [z_{10}; z_{20}, z_{21}, z_{22}, z_{2k}; f](z_2 - z_{20})(z_2 - z_{21})(z_2 - z_{22}) + \\
& + [z_{10}; z_{20}, z_{21}, z_{22}, z_{2k}, z_{2(k+1)}; f](z_2 - z_{20})(z_2 - z_{21})(z_2 - z_{22})(z_1 - z_{2k}).
\end{aligned}$$

Ugyanekkor a 4.2. Tétel (ii) részének segítségével az alábbi, felülről közelítő, interpolációs polinomot kapjuk:

$$\begin{aligned}
(4.10) \quad L_I(z_1, z_2) = & \\
& = [z_{1n_1}; z_{2n_2}; f] + [z_{1n_1}, z_{1(n_1-1)}; z_{2n_2}; f](z_1 - z_{1n_1}) + \\
& + [z_{1n_1}; z_{2n_2}, z_{2(n_2-1)}; f](z_2 - z_{2n_2}) + \\
& + [z_{1n_1}, z_{1(n_1-1)}; z_{20}, z_{2(n_2-1)}; f](z_1 - z_{1n_1})(z_2 - z_{2n_2}) + \\
& + [z_{1n_1}, z_{1(n_1-1)}, z_{1(n_1-2)}; z_{2n_2}; f](z_1 - z_{1n_1})(z_1 - z_{1(n_1-1)}) + \\
& + [z_{1n_1}, z_{1(n_1-1)}, z_{1(n_1-2)}, z_{1i}; z_{2n_2}; f] \times \\
& \times (z_1 - z_{1n_1})(z_1 - z_{1(n_1-1)})(z_1 - z_{1(n_1-2)}) + \\
& + [z_{1n_1}, z_{1(n_1-1)}, z_{1(n_1-2)}, z_{1i}, z_{1(i+1)}; z_{2n_2}; f] \times \\
& \times (z_1 - z_{1n_1})(z_1 - z_{1(n_1-1)})(z_1 - z_{1(n_1-2)})(z_1 - z_{1i}) + \\
& + [z_{1n_1}; z_{2n_2}, z_{2(n_2-1)}, z_{2(n_2-2)}; f](z_2 - z_{2n_2})(z_2 - z_{2(n_2-1)}) + \\
& + [z_{1n_1}; z_{2n_2}, z_{2(n_2-1)}, z_{2(n_2-2)}, z_{2i}; f] \times \\
& \times (z_2 - z_{2n_2})(z_2 - z_{2(n_2-1)})(z_2 - z_{2(n_2-2)}) + \\
& + [z_{1n_1}; z_{2n_2}, z_{2(n_2-1)}, z_{2(n_2-2)}, z_{2k}, z_{2(k+1)}; f] \times \\
& \times (z_2 - z_{2n_2})(z_2 - z_{2(n_2-1)})(z_2 - z_{2(n_2-2)})(z_2 - z_{2k}).
\end{aligned}$$

Ha a (4.9), (4.10) polinomokban a  $z_1, z_2$  változókat az  $X_1, X_2$  valószínűségi változókkal helyettesítjük, és vesszük a nyert valószínűségi változók várható értékét, majd azokat a 4.1. és a 4.2. Tételekben felhasználjuk, akkor éles alsó és felső korlátokat kapunk  $E[f(X_1, X_2)]$  értékére. Az  $X_1, X_2$  valószínűségi változók fent említett polinomjai várható értékét könnyen megkaphatjuk, ha ismerjük az

$$E(X_j^k), \quad k = 1, 2, 3, 4, \quad j = 1, 2$$

momentumokat és a  $\text{Cov}(X_1, X_2)$  mennyiséget.

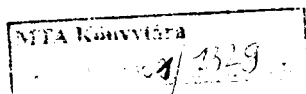
A témában elmélyülni szándékozó olvasó számára az eddig idézetteken kívül ajánljuk még Nagy és Prékopa [4] cikkét.

### Irodalom

- [1] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press (New York, 1991).
- [2] C. Jordan, *Calculus of Finite Differences*, Chelsea Publishing Company (New York, 1947).
- [3] C. E. Lemke, The Dual Method for Solving the Linear Programming Problem, *Naval Research Logistics Quarterly*, 1 (1954), 36–47.
- [4] G. Nagy and A. Prékopa, *On Multivariate Discrete Moment Problems and their Applications to Bounding Expectations and Probabilities*, RUTCOR Research Report (2000), 41–2000.
- [5] T. Popoviciu, Les Fonctions Convexes, *Actualités Scientifiques et Industrielles*, 992 (1944), Hermann (Paris).
- [6] A. Prékopa, Inequalities on Expectations Based on the Knowledge of Multivariate Moments, in: *Stochastic Inequalities*, Ed. M. Shaked and Y. L. Tong, Institute of Mathematical Statistics, Lecture Notes — Monograph Series, Vol 22 (1992), pp. 309–331.
- [7] A. Prékopa, A Brief Introduction to Linear Programming, *The Mathematical Scientist*, 21 (1996), 85–111.
- [8] A. Prékopa, Bounds on Probabilities and Expectations Using Multivariate Moments of Discrete Distributions, *Studia Scientiarum Mathematicarum Hungarica*, 34 (1998), 349–378.
- [9] A. Prékopa, *On Multivariate Discrete Higher Order Convex Functions and their Applications*, RUTCOR Research Report (2000) 39–2000. Also in: Proceedings of the Sixth International Conference on Generalized Convexity and Monotonicity, Karlovasi, Samos, Greece, August 29–September 2, to appear.

(Beérkezett: 2001. február 13.)

NAGY GERGELY  
BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM  
MATEMATIKA INTÉZET  
DIFFERENCIÁLEGYENLETEK TANSZÉK  
1111 BUDAPEST, MŰEGYETEM RAKPART 1-3.  
E-mail: GNAGY@MATH.BME.HU





PRÉKOPA ANDRÁS  
ELTE TERMÉSZETTUDOMÁNYI KAR  
OPERÁCIÓKUTATÁSI TANSZÉK  
1053 BUDAPEST, KECSKEMÉTI U. 10-12.  
*E-mail:* PREKOPA@CS.ELTE.HU  
RUTCOR, RUTGERS CENTER FOR OPERATIONS RESEARCH  
RUTGERS UNIVERSITY  
640 BARTHOLOMEW ROAD, PISCATAWAY, NJ 08854-8003  
*E-mail:* PREKOPA@RUTCOR.RUTGERS.EDU

## ONE-SIDED APPROXIMATION OF MULTIVARIATE DISCRETE FUNCTIONS BY POLYNOMIALS

GERGELY NAGY AND ANDRÁS PRÉKOPA

The main result of the paper is a formula for the difference between a multivariate function and a corresponding Lagrange interpolation polynomial. It generalizes the well-known univariate formula, where the difference between the function and the unique Lagrange polynomial is equal to the product of a higher order divided difference of the function and differences of the variable and the base points. By suitable choices of the base points of the multivariate polynomial we can ensure the nonnegativity or nonpositivity of the residual term and create one-sided approximations to the function. If we replace random variables for the deterministic ones and take expectations, we can obtain bounds on the given function of the random variables involved random variables, based on moment information.

A kiadásért felelős a BJMT főtítkára  
Szedte és tördelte az Egyenes Bt.  
Nyomta az MSZH Nyomda és Kiadó Kft., Budapest  
Felelős vezető: Nagy László  
Budapest, 2000  
Megjelent 10 (A/5) ív terjedelemben  
250 példányban  
HU ISSN 0133-3399

**MAGYAR  
TUDOMÁNYOS AKADÉMIA  
KÖNYVTÁRA**

## ÚTMUTATÁS A SZERZŐKNEK

Az Alkalmazott Matematikai Lapok csak magyar nyelvű dolgozatokat közöl. A kéziratok gépelését olyan formában kérjük, hogy minden gépelt oldal 25, egyenként átlag 50 betűhelyes sort tartsalmazzon. A közlésre szánt dolgozatokat három példányban kell beküldeni. Előnyben részesülnek a  $\text{\TeX}$ -ben elkészített dolgozatok. Ezeket két kinyomtatott példány kíséretében diszketten kérjük beadni.

A kéziratok szerkezeti felépítésének a következő követelményeket kell kielégíteni. A fejlécnek tartalmaznia kell a dolgozat címét, a szerző teljes nevét, valamint annak a városnak a nevét, ahol a szerző dolgozik. A fejléc után egy, képletet nem tartalmazó, legfeljebb 200 szóból álló kivonatot kell minden esetben megadni. A dolgozatot címmel ellátott szakaszokra kell bontani, és az egyes szakaszokat arab sorszámozással kell ellátni. Az esetleges bevezetésnek mindig az első szakaszt kell alkotnia. Az irodalomjegyzék után, a kézirat befejezésekképpen fel kell tüntetni a szerző teljes nevét és a munkahelye (illetve lakása) pontos címét. A dolgozatban előforduló képleteket szakaszonként újrakezdődően, a képlet előtt két zárójel közé írt kettős számozással kell azonosítani. Természetesen nem szükséges minden képletet számozással ellátni. Az esetleges definíciókat és tételeket (segédttételeket és lemmákat) ugyancsak szakaszonként újrakezdődő, kettős számozással kell ellátni. Kérjük a szerzőket, hogy ezeket, valamint a tételek bizonyítását a szövegben kellő módon emeljék ki. Minden dolgozathoz csatolni kell egy angol, német francia vagy orosz nyelvű, külön oldalra gépelt összefoglalót. Amennyiben lehetséges, kérjük a nyomtatás számára különösen nehézkes matematikai jelölések használatának az elkerülését.

A dolgozatok ábráit és az esetleges lábjegyzeteket a dolgozat végén, különálló lapokon kérjük beküldeni. Mind az ábrákat, mind a lábjegyzeteket a dolgozat szakaszokra bontásától függetlenül, folytatólagos arab sorszámozással kell ellátni. Az ábrák elhelyezését a dolgozat megfelelő helyén, széljegyzetként feltüntetett, ábraazonosító sorszámokkal kell megadni. A lábjegyzetekre a dolgozaton belül az azonosító sorszám felső indexkénti használatával lehet hivatkozni.

Az irodalmi hivatkozások formája a következő. Minden hivatkozást fel kell sorolni a dolgozat végén található irodalomjegyzékben, a szerzők, illetve a társszerzők esetén az első szerző neve szerint alfabetikus sorrendben úgy, hogy a cirill betűs szerzők nevét a Mathematical Reviews átirási szabályai szerint latin betűsre kell átírni. A folyóiratban megjelent cikkekre [1], a könyvekre [5], a kötetben megjelent dolgozatokra [4], a disszertációkra [3] és a gépi program leírásokra [2] a következő minta szerint kell hivatkozni:

- [1] Farkas, J., Über die Theorie der einfachen Ungleichungen, *Journal für die reine und angewandte Mathematik* 124 (1902) 1–27.
- [2] Kéri, G., „DUALSIMP”, rutin a CDC 3300-ás gépekre (Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézete, CDC 3300 felhasználói ismertetők 2. 1973. május) 19–20.
- [3] Prékopa, A. „Sztochasztikus rendszerek optimalizálási problémáiról”, doktori értekezés. Magyar Tudományos Akadémia, Budapest, 1970.
- [4] Prabhu, N. U. „Recent research on the ruin problem of collective risk theory”, in: *Inventory Control and Water Storage* Ed. A. Prékopa (János Bolyai Mathematical Society and North-Holland Publishing Company, Amsterdam–London, (1973) 221–228.
- [5] Zoutendijk, G. *Methods of Feasible Directions* (Elsevier Publishing Company, Amsterdam and New York, 1960).

A dolgozatok szövegében az irodalmi hivatkozás számait szögletes zárójelben kell megadni, mint például [5] vagy [4, 76–78]. A szerzők a dolgozatukról 50 darab ingyenes különlenyomatot kapnak. A dolgozatok után szerzői díjat az Alkalmazott Matematikai Lapok nem fizet.

## TARTALOMJEGYZÉK

<i>Schusztér György</i> , Fuzzy döntési eljárások összehasonlítása irányítástechnikai szempontból .	113
<i>Kászonyi László</i> , Bináris kódok generálása környezetfüggetlen nyelvtanok segítségével . . . . .	133
<i>Faragó István és Kovács Mihály</i> , Ergodikus Markov-láncok és alkalmazásaik . . . . .	149
<i>Fazekas Attila és Sánta István</i> , Vázkijelölés kontúr-reprezentációjú képeken . . . . .	165
<i>Bognár Katalin</i> , Leíró logikák az ismeretábrázolásban . . . . .	183
<i>Nagy Gergely és Prékopa András</i> , Függvények féloldalas approximációja polinomokkal . . . . .	195

## INDEX

<i>György Schusztér</i> , Comparison of Fuzzy reasoning methods for process control . . . . .	113
<i>László Kászonyi</i> , How to generate binary codes using context-free grammars . . . . .	133
<i>István Faragó and Mihály Kovács</i> , Ergodic Markov chains and their applications . . . . .	149
<i>Attila Fazekas and István Sánta</i> , Skeletonization on contour-representation . . . . .	165
<i>Katalin Bognár</i> , Description logics in knowledge representation . . . . .	183
<i>Gergely Nagy and András Prékopa</i> , One-sided approximation of multivariate discrete functions by polynomials . . . . .	195